Basketball Shooting as a Model System for Understanding Skill Learning and Motor Variability

HARVARD UNIVERSITY

Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Harvard John A. Paulson School of Engineering and Applied Sciences
have examined a dissertation entitled:

"Basketball Shooting as a Model System for Understanding Skill Learning and Motor Variability"

presented by: Rishi Bal Singh

Signature _____

Typed name: Professor M. Smith

Signature _____

Typed name: Professor R. Howe

Signature _____

Typed name: Professor B. Ölveczky

May 2, 2022

Basketball Shooting as a Model System for Understanding Skill Learning and Motor Variability

A dissertation presented

by

Rishi Bal Singh

to

The School of Engineering and Applied Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Engineering Sciences

Harvard University

Cambridge, Massachusetts

May, 2022

Dissertation Advisor: Maurice A. Smith Rishi Bal Singh

Basketball Shooting as a Model System for Understanding Skill Learning and Motor Variability

ABSTRACT

The motor system has the incredibly ability to learn complex movements which require fine control and coordinated timing of numerous joints and limbs. The study of motor learning and motor control has traditionally studied simpler movements in order to control the complexity well enough to have resolution on particular subsystems and components of interest. Virtual tasks and virtual systems have been incredibly helpful by allowing precise measurement of movement and precise control over movement and feedback variables. However, little is known about how well the knowledge gained of the components of motor control generalizes to complex, real physical tasks.

The first part of this work deals with designing a system capable of precisely measuring movement and performance in real basketball free-throw shooting. Free-throw shooting is particularly well-positioned as a model system in which to study complex motor skills because of its naturally controlled nature, its widespread popularity, and as a complex throwing motion. In particular, the system is focused towards understanding the role of longer feedback latencies in motor skill learning and performance, which are hypothesized to impair components of motor learning crucial to complex motor skills. Overall, the system is capable of predicting shot outcomes with high accuracy and providing shooters with predictive feedback about their shot's future outcome in just 37 ms, reducing the feedback delay in basketball shooting by a factor 25x.

The second part of this work investigates the use of this system as a training tool for shooters. Motor learning studies on arm-reaching suggest that delays in feedback specifically impair implicit learning,

which is thought to be crucial for complex skill learning. However, these studies have mostly focused on shorter latencies of hundreds of milliseconds. Throwing tasks naturally contain longer delays on the order of seconds, and it is unclear whether these tasks are subject to similar decrements in performance from their inherent latencies. For example, for a free-throw shot it takes the ball ~1 second after it leaves hand to reach the rim. Interestingly, data indicates that even among professionals free-throw performance is mediocre, suggesting that there may be factors limiting performance in the task. We trained shooters in two groups, one which received predictive feedback of their shot outcome just 37 ms after releasing the ball, and another which received the same predictive feedback but delivered concurrent with the ball arriving at the rim. Shooters were trained over 6 days of shooting with 3200 total shots. Neither group displayed detectable changes to overall performance over, though upon dissection of the ball's landing location we did see an improvement in alignment of mean shot landing location with an optimal success region for shooters receiving advanced feedback, suggesting the potential for time-advanced feedback to improve motor skill performance in improved experimental designs.

In the last section of this work we use the basketball tracking and prediction system's capability for precision measurement of basketball shot outcome to investigate whether transient fluctuations in performance variability can be detected among sequences of free-throw outcomes. This idea has been called the "hot hand", which in basketball refers to a player who seems to be temporarily shooting noticeably above expectation. However, in 1985 an influential study staunchly claimed, despite widespread belief in the "hot hand", that it was a cognitive fallacy, and that such streaks in performance were no different than what could be expected by random chance. Since this study, existence of performance fluctuations in basketball shooting has been a hotly contested, though a lack of statistical power in the methods and datasets used thus far have limited the strength of findings on both sides of the debate. We leverage a method which gauges the ability to predict future performance in shot sequences, and combine this method with precise continuous-valued measurements of shot accuracy. Together, this drastically decreases the number of shots needed to detect significant temporal fluctuations in performance by 10-14x over existing methods. Using this new method on the largest controlled shooting

dataset collected to date, we decisively detect the presence of performance fluctuations in individual shooters, proving that performance fluctuations are actually a widespread phenomenon across shooters.

TABLE OF CONTENTS

Contents

Acknowledgements

I knew doing a PhD was going to be a marathon and not a sprint. Few things, however, prepared me for the whirlwind that has been the past few months as I prepare my thesis, and I find myself with little emotional breadth to truly thank those who have helped me along my way with words fitting of the love and support they have given me. Here, I have attempted to encapsulate my appreciation for these many people, though my words will certainly fall short of their impact on my time at Harvard and beyond.

Foremost, I would like to thank my parents. They taught me to chase after my dreams, and to always light a flame for those you love before you go on your way.

Special thanks are due to my two older sisters, who have powerfully shaped my who I am today. Rashmi, I always thought you were the coolest and I wanted to grow up to be just like you. Renu, you taught me to find the hidden details in everything around me. And you introduced me to basketball. There is no doubt that the love and care I received from both of you has afforded me the successes I have today.

To my labmates over the many years, the time I have been able to spend with you all has been a beautiful and unexpected gift. You have given me laughs, friendship, love, support, and guidance through the many ups and downs I experienced over the past 7 years. It goes without saying that I would not have made it this far without every one of you, and so thank you Yohsuke, Andrew, Ryan, Laith, Tanvi, George, Hayoun, Andrew, and Sarah.

To my adopted brothers Poh, Arash, Adam, Samer, and George, thank you for being there for me with open doors and open arms. I look up to every one of you, and you have kept me moving forward when times were difficult.

To Will, thank you for your steady and warm friendship, and thank you for weathering the daily onslaught of complaints about lab work and my project.

Finally, I need to thank Joe Pemberton. Hardly a day goes by when I don't think of our times together, and the summers we spent working out have been a source of motivation and strength for me. I was obsessed with movement before I met you, but you taught me to turn movement into vocabulary, and I would not be the person I am today without your friendship.

Introduction

For as long as I can remember, movement has meant freedom. For a period of time it quite literally meant freedom, as I had not yet developed the ability to escape when my sister trapped me in the laundry basket. Lucky for me I am not claustrophobic, but for many the feeling of being constrained in their ability to move, even when all they intend is to simply stand still in an elevator, is incredibly uncomfortable; for young kids, having to sit still in a chair for more than twenty minutes can pose a real challenge; one could even argue that freedom of speech is, in a way, also freedom of movement, as speech requires movement of air and skilled control of the muscles of your mouth to produce words.

In another way, separate from speaking, our movements can belie what is happening internally: our friends can tell when we are happy by the pep in our step, and when we are tired by the slouch in our posture. Maybe it is better to say that freedom of movement is closer to freedom of expression. Undeniably though, movement is well-woven into the fabric of most moments in our life. It is humbling then, to realize how little we understand about it. As an example, in 2017 a video from Boston Dynamics of their state-of-the-art bipedal robot Atlas went viral for doing something most kids can do as toddlers – jumping over boxes. This is not to say that the engineers and developers accomplished anything short of amazing, and I was just as astonished as others were watching a humanoid robot jump over boxes, but it highlights the complexity of even simple movements that humans accomplish with ease.

In general, studying the types of skilled motions we do every day, like tossing paper into the trash, is challenging because they often combine several different modalities that we do not fully understand even in isolation. So far, we have resorted to studying pieces of the puzzle by breaking more complex movements down into simpler, more controllable movements that can be studied in a lab setting. An important facilitator for such studies has been the rise of human-computer interfaces (HCIs) which allow experimenters to develop virtual tasks in which they can control several factors which might affect a motor skill such as visual feedback or the mapping between real motions of a person's limbs and corresponding motion on a display (Izawa et al., 2002; Krakauer et al., 1999; Shadmehr & Mussa-Lvaldi, 1994). By altering a person's sensory experience during the task, for example making the screen cursor

move left when your hand moves right, and measuring the resulting movements, researchers can try to infer details about the internal processes occurring in the brain and the motor system. Though the movements involved are often simple point-to-point arm-reaching movements, they have revealed several core features of the motor system, one of which is the existence of two general categories of motor learning: explicit learning, which refers to strategic high-level thinking about the desired movement and movement goal, and implicit learning, which refers to the mostly subconscious process responsible for complicated adjustments to multiple joints, limbs and muscles (Mazzoni & Krakauer, 2006). As the complexity of the motor skill increases, generally meaning that it involves more limbs and multiple movement goals, the number of types of adjustments one could make expands exponentially, and so it is thought that the largely subconscious implicit learning is crucial to learning more complex motor skills.

These virtual tasks have continued to prove fertile for understanding the nuances of motor learning. However, the HCI systems they depend on are, by nature, subject to inherent latencies between a person's movement and the resulting movement shown on displays. These latencies occur because of the time needed to detect the movement, process the movement, and make the feedback available either visually or through another medium. Though these feedback delays are usually only tens to hundreds of milliseconds, a growing body of work suggests that they can impair motor performance and motor learning processes. Across a variety of applications, such as surgical training, steering, and virtual reality systems, HCI system latencies have been found to cause decrements in human performance and poorly affect user experience (Friston et al., 2016; Kaber et al., 2012; Meehan et al., 2003). For example, a study of robotic surgeries among expert surgeons reported decrements in surgical performance with latencies as small as 70 ms, and a nearly 60% increase in surgery completion time for latencies of 160 ms (Kumcu et al., 2017).

Behavioral studies in both humans and animals suggest that small latencies can specifically impair motor learning processes. In classical conditioning, trace conditioning paradigms have studied the effect of latencies in animals by introducing delays between an unconditioned stimulus (US) and a conditioned

stimulus (CS). In this simple form of associative learning where an animal learns to perform a movement in response to a signal, for example closing their eyelid to avoid an airpuff (US) which consistently follows presentation of a tone (CS), added latencies between CS and US presentation have been shown to generally impair learning, with 400 ms delays being reported to slow learning by 50% (Kehoe et al., 2009; Smith et al., 1969). In human motor learning studies of arm-reaching movements, the effects of latencies have been studied by artificially increasing the latency between measured and displayed hand motion in HCI setups. Here, subjects learn to alter their hand paths to compensate for rotational shifts in the virtual hand paths shown on a display, which they are mostly able to do even in the presence of display latencies. However, the contribution of implicit learning in the presence of display latencies is markedly decreased, requiring subjects to instead rely on higher level thinking about the task to succeed in the movements (Brudner et al., 2016; Schween & Hegele, 2017). In the case of a display rotation, this amounts to strategically aiming left or right depending on what direction you are expecting the rotation to be. This suggests that the negative impact of feedback latencies could be amplified as movements become more complex, as implicit learning is thought to be crucial for learning more complex motor skills.

It is interesting then to consider skilled throwing motions, which are complex multi-joint movements that often have much long feedback delays on the order of seconds. This is caused by the physical time it takes an object to reach its target after being released. Despite this, people display incredible skill and accuracy in throwing motions, like a professional quarterback in the NFL or an outfielder who can throw a runner out at home base. In fact, beyond its place in modern sports, throwing motions have been posited to be a core feature of human evolution (Roach et al., 2013; Young, 2003). Together, it is unclear what the role of feedback latency is in such complex movements and what factors control learning and acquisition of these skills, and so the goal of the presented work is to better understand the role of feedback latency in complex motor skills like throwing.

Historically, study of complex motor skill has resorted to simplified or modified movements that allow for easier experimental manipulation and control. The use of virtual tasks such as those used in arm-

reaching studies has allowed researchers very fine control over several important factors in motor learning, such as visual and proprioceptive feedback, which has been crucial in achieving a greater understanding of many nuanced components of the motor system. However, there remains a large gap between the movements studied in lab settings and skills like throwing or dancing, partly because complex movements are more difficult to manipulate and can have more factors to control. Recent work suggests that even detailed virtual recreations of a task like throwing lack the richness of real physical environments, which can affect how well conclusions from studies using such virtual tasks generalize to the reality of complex motor skills (Zhang & Sternad, 2021). Thus, there is a need for systems which afford experimental control in complex skills without removing them from their natural, physical environment.

We aimed to thus build a system that affords control over feedback and feedback timing in a real, physical complex task, as opposed to a virtual task, with minimal intervention. Free-throw shooting is well positioned as a model skill for such a system as it is unique among sports in its controlled nature, and because the widespread popularity of the sport allows easy access to experienced shooters with years of practice. The first part of this work is focused on developing a system for studying basketball free-throw shooting with minimal intervention, with the particular focus and motivation of implementing a predictive delay cancellation system for manipulating the feedback latency naturally occurring in basketball shooting. While most previous work has studied the effect of adding latencies to movements which usually have no feedback delay, we instead choose a movement with inherent natural delay and study its effect by circumventing it. This requires a predictive system that can predict the shot outcome based on early information during the shooting motion and then make this predictive feedback available to the shooter provide well before the ball reaches the rim.

In the second part of this work we use the developed system in an experimental design to study the effect of essentially removing feedback latencies in free-throw shooting. We do this by training participants in two groups: one group which receives feedback about the outcomes of their shots just 35ms after they

release the ball, and another group which receives this same type of feedback though the feedback is only available concurrent to the ball passing through the rim.

In the third part of this work, we use the available shot data we collected during development of the system and during experimentation on feedback latencies to investigate a novel method for detecting performance fluctuations in individual shooters. For basketball shooting, this amounts to temporal changes in a shooter's inherent variability when performing the shooting motion. The existence of such performance fluctuation, generally known as the "hot hand", has been debated since a landmark paper in 1985 claimed that it was a cognitive fallacy (Gilovich et al., 1985). We introduce a new method which, combined with our system, provides tremendously increased statistical power over existing methods. We use this method to show, for the first time, strong definitive evidence that the hot hand exists in individual shooters.

References for introduction

Brudner, S. N., Kethidi, N., Graeupner, D., Ivry, R. B., Taylor, J. A., & Taylor, J. A. (2016). Delayed feedback during sensorimotor learning selectively disrupts adaptation but not strategy use. *J Neurophysiology*, *115*, 1499–1511. https://doi.org/10.1152/jn.00066.2015.-In

Friston, S., Karlström, P., & Steed, A. (2016). The effects of low latency on pointing and steering tasks. *IEEE Transactions on Visualization and Computer Graphics*, *22*(5), 1605–1615. https://doi.org/10.1109/TVCG.2015.2446467

Gilovich, T., Vallone, R., & Tversky, A. (1985). The Hot Hand in Basketball: On the Misperception of Random Sequences. In *COGNITIVE PSYCHOLOGY* (Vol. 17).

Izawa, J., Kondo, T., & Ito, K. (2002). *Biological Robot Arm Motion th.rough Reinforcement Learning*.

Kaber, D. B., Li, Y., Clamann, M., & Lee, Y. S. (2012). Investigating human performance in a virtual reality haptic simulator as influenced by fidelity and system latency. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, *42*(6), 1562–1566. https://doi.org/10.1109/TSMCA.2012.2201466

Kehoe, E. J., Ludvig, E. A., & Sutton, R. S. (2009). Magnitude and Timing of Conditioned Responses in Delay and Trace Classical Conditioning of the Nictitating Membrane Response of the Rabbit (Oryctolagus cuniculus). *Behavioral Neuroscience*, *123*(5), 1095–1101. https://doi.org/10.1037/a0017112

Krakauer, J. W., Ghilardi, M.-F., & Ghez, C. (1999). Independent learning of internal models for kinematic and dynamic control of reaching. *Nature Neuroscience*, *2*(11), 1026–1031.

Kumcu, A., Vermeulen, L., Elprama, S. A., Duysburgh, P., Platiša, L., van Nieuwenhove, Y., van de Winkel, N., Jacobs, A., van Looy, J., & Philips, W. (2017). Effect of video lag on laparoscopic

surgery: correlation between performance and usability at low latencies. *International Journal of Medical Robotics and Computer Assisted Surgery*, *13*(2). https://doi.org/10.1002/rcs.1758

Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *Journal of Neuroscience*, *26*(14), 3642–3645. https://doi.org/10.1523/JNEUROSCI.5317-05.2006

Meehan, M., Razzaque, S., Whitton, M. C., & Brooks, F. P. (2003). *Effect of Latency on Presence in Stressful Virtual Environments*.

Roach, N. T., Venkadesan, M., Rainbow, M. J., & Lieberman, D. E. (2013). Elastic energy storage in the shoulder and the evolution of high-speed throwing in Homo. *Nature*, *498*(7455), 483–486. https://doi.org/10.1038/nature12267

Schween, R., & Hegele, M. (2017). Feedback delay attenuates implicit but facilitates explicit adjustments to a visuomotor rotation. *Neurobiology of Learning and Memory*, *140*, 124–133. https://doi.org/10.1016/j.nlm.2017.02.015

Shadmehr, R., & Mussa-Lvaldi, F. A. (1994). *Adaptive Representation of Dynamics during Learning of a Motor Task*.

Smith, M. C., Coleman, S. R., & Gormezano, I. (1969). Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous, and forward CS-US intervals. *Journal of Comparatvie and Physiological Psychology*, *69*(2), 226–231.

Young, R. W. (2003). Evolution of the human hand: the role of throwing and clubbing. In *J. Anat* (Vol. 202).

Zhang, Z., & Sternad, D. (2021). Back to reality: Differences in learning strategy in a simplified virtual and a real throwing task. *Journal of Neurophysiology*, *125*(1), 43–62. https://doi.org/10.1152/JN.00197.2020

Chapter 1: System design for predictive delay cancellation in basketball free-throw shooting

1.1 Abstract

Towards studying the effects of feedback latency in basketball free-throw shooting, we aimed to design and build a predictive system that can circumvent the naturally occurring ~1 s wait time between shot release and the ball's arrival at the rim. Predicting the shot outcome based on information available early during the shot would allow such a system to provide feedback before the ball reaches the rim, reducing the inherent feedback delay. While body kinematics provide the earliest available information in the shot, they have unfortunately led to poor predictive performance, which might be attributed to idiosyncrasies across shooters. Hence we instead focus on predicting shot outcome based on initial motion of the ball, because after release, its flight-path is completely determined by projectile physics. We successfully built such a system using a custom stereo-camera setup to track the ball's initial motion state, which avoids the use of less durable markers and stickers that commercial motion tracking systems require. We then use the tracked initial linear and rotational motion of the ball in a predictive model to determine where a shot will land relative to the rim, and then use this shot endpoint as input to a classification model to ultimately predict binary success of the shot.

With 40 μm precision in tracking ball position and 0.08 rev/s precision in tracking ball spin, our system is able to predict shot landing location within 2.5 cm using only 12.5ms of tracked ball flight. This is a marked improvement over previous work which, at best, reported 4 cm precision while fitting a model using the full extent of the ball's flight path (Nakano et al., 2020). This would likely equate to ~8 cm of precision had the authors predicted using only initial ball motion as done in our system, and so we estimate that linear units our system is a ~70% improvement over previous work.

Predicted shot endpoints are fed into a classification model based on a decontaminated definition of shot success which predicts binary shot success with 99% accuracy. This highly accurate feedback can be made available to shooters within just 35 ms after they release the ball, constituting a ~25x factor reduction in feedback latency.

1.2 Introduction

We are interested in studying the effect of long feedback latencies on motor skill performance, for which basketball is well-positioned as a model system. Previous work on feedback latency has been limited to simpler virtual movements where latencies on the order of hundreds of milliseconds are added to movements which normally have no delays in feedback. While these latencies are very relevant to the types of latencies found in virtual tasks using human-computer interface (HCI) systems, they are in stark contrast to the latencies commonly seen in throwing motions, which are ubiquitous across sports and even posited to have played a role in human evolution (Brudner et al., 2016; Roach et al., 2013; Schween & Hegele, 2017; Young, 2003). We thus aim to study feedback latency in a throwing task, specifically basketball free-throw shooting, where manipulation of feedback latency requires predictive *elimination* of the existing physical delay incurred by travel time of the ball through the air.

There has been limited work on predicting basketball shot outcomes from early information in the shot. One possible approach is to use kinematic information about the shooter's body as they go through the shooting motion, which is advantageous because it is the earliest available information. Some studies have attempted to use such measurements during the shooting motion to predict whether the ball will pass through the hoop, though accuracy in their predictions has been weak (Ammar et al., 2016; Malone et al., 2002; Mullineaux & Uhl, 2010; Özkan et al., 2019).

Another approach to predicting shot outcome is to use the initial motion state of the ball measured immediately after it leaves the shooter's hands. This is a more viable approach, as ball motion once it leaves the shooter's hands is completely determined by physics. This will be consistent across shooters, as opposed to body kinematics which are idiosyncratic. Additionally, while there is already much existing knowledge from classical mechanics about projectile motion, there is also existing work specifically on the physics of free-throw shots that we can leverage (Silverberg et al., 2003; Tran & Silverberg,

2008).Thus, we aim to predict shot outcomes using measurements of the ball's initial motion state as soon as is possible after the ball leaves the shooter's hands.

Here I outline the design of a system which uses rich visual information from cameras to manage a number of tasks relevant to performing experiments in basketball shooting; most importantly, however, we use these cameras to obtain very precise estimates of the ball's initial motion state. This system is comprised of tracking algorithms which detect the linear and rotational motion of the ball, a classification model based on a novel decontaminated measure of binary shot success, and a predictive model which uses early information of the ball's motion to predict it's landing location at the rim.

1.3 Camera-based shot tracking

The main goal of our system is to obtain highly precise estimates of the ball's initial state, including its position, linear motion, and rotational motion soon after release in order to predict shot outcome. However, the system also needs to be able to detect and track other events such as whether the shot passes through the rim, the trajectory of its approach at the rim, and how it reacts after contact with the rim and backboard to provide the requisite training data for designing and fitting models to predict shot outcome. Further, the system would need to accurately detect events such as release of the ball from the shooter's hand and accurately ignore benign events such as motion from passes between shots to facilitate uninterrupted shooting during experiment sessions. Visual information is an immensely rich source of information which can encode several attributes such as motion, state, and appearance of a wide variety of events and objects, and so we choose to use cameras as the tracking elements our system to leverage this richness and the afforded flexibility to track various events of interest throughout an experiment.

Stereo vision systems, which combine simultaneously captured visual information across two or more cameras to estimate 3D position of detected objects, have been widely used for a variety of precision tracking systems (Tippetts et al., 2016). A typical free-throw shot occurs over a large volume of space, as the vertical position of the ball can vary anywhere from 6 ft above the ground when it is released by the

shooter, up to peak heights of 14 feet or more before descending towards the rim; longitudinally, the ball traverses a distance over 12 ft. As such, we decided to use wide-angle lenses (12mm HP Series Lens, 1", Edmund Optics) for our cameras, which maximizes the trackable volume of our system. Additionally, minimizing the system's prediction latency requires cameras capable of high frame rates and high data transfer rates, though faster data transfer is often facilitated by lower image resolutions, and so comes at the cost of measurement precision. For our system we used 2.3-megapixel USB 3.1 machine vision cameras capable of framerates up to 163 frames/sec (Grasshopper 3 GS3-U3-23S6C, Teledyne FLIR), which reasonably meets our design requirements. Fig. 1.1 shows an annotated picture of the room in which the system is installed. We installed two cameras, one above and one lateral to the shooter, to track the ball during shot release; two more cameras, one above and one lateral to the rim, were installed to track the ball's trajectory as it approached the rim, detect how it bounced if contact with the rim was made, and to detect whether the shot was a make or a miss. Image capture across cameras was synchronized using an external signal generator connected to pins controlling each cameras' trigger.

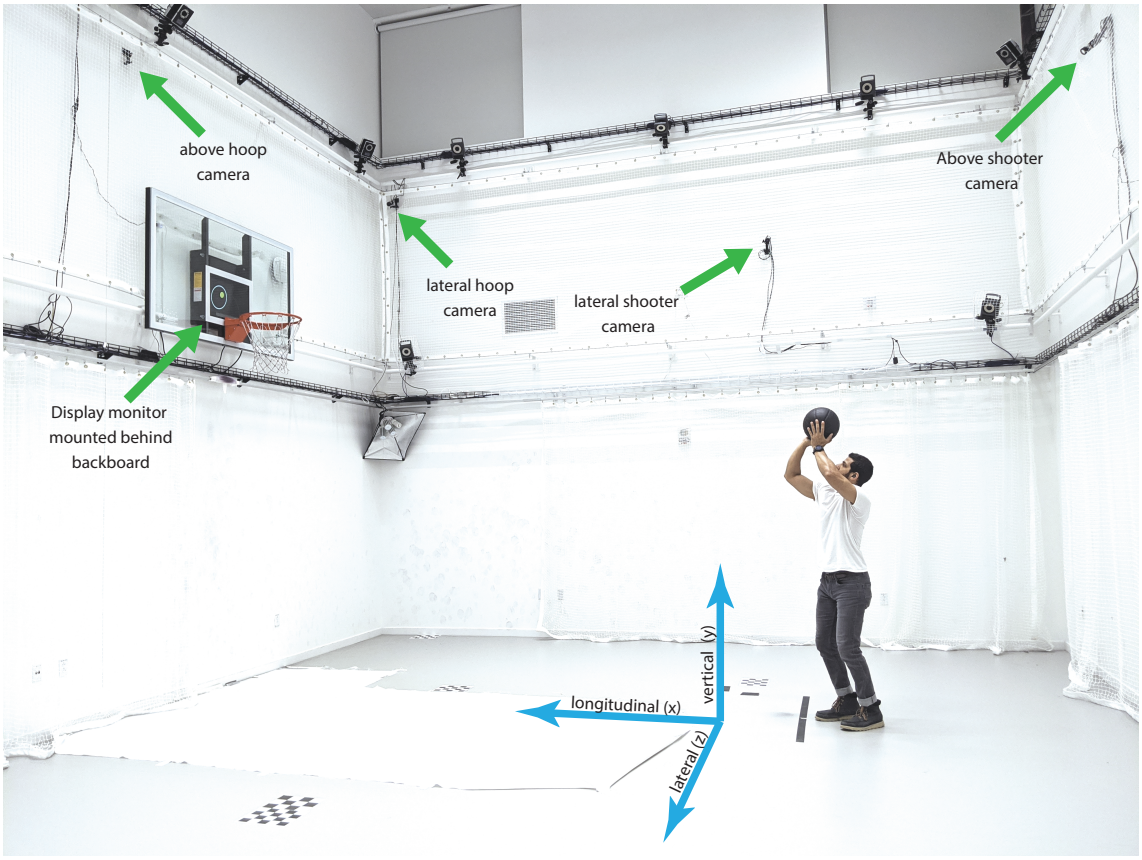

Figure 1.1: Image of installed system components in experimental shooting room

The system consists of two zones monitored by separate pairs of synchronized cameras running at 80 frames/sec (2.3 MP, USB3.1). An above shooter camera and lateral shooter camera monitor the initial release and motion of the ball, while an above hoop and lateral hoop monitor the ball from as it arrives at the rim until it falls to the ground. Images from pairs of cameras are sent in real-time to a computer which performs 2D ball tracking in individual images from each camera. This 2D information is combined across cameras using precisely calibrated camera parameters to produce 3D estimates of ball position, linear velocity, and rotational motion. A monitor mounted behind the glass backboard is positioned to be in the shooter's line of sight as they shoot, facilitating delivery of visual feedback to the shooter during or after the shot. The convention and naming of coordinate axes aligned to the direction of the shot are shown for reference. The longitudinal (x) direction refers to the direction pointing towards the rim from the shooter. The lateral (z) direction encodes left to right from the shooter's point of view. The vertical (y) direction is the axis aligned with the direction of gravity.

1.3.1 Calibration for image distortion correction and stereo camera measurement

Using wide-angle lenses allows our system to monitor events across a large space by directing a wider swath of incoming light towards the sensor inside the camera. A consequence of increasing the viewing

range with the lens however is warping of the incoming image as the spherical lens bends the incoming light. Additionally, imprecision in perpendicular alignment of the camera light sensor relative to the direction of incoming light during manufacturing can also introduce warping of the image. Making accurate measurements of any detected object's position in space requires precise correction of these image distortions. Following commonly used procedures which use images of chessboard grid patterns at several viewing angles, we calibrated parameters in a model describing image distortion arising from spherical lenses and lens-sensor misalignment for each individual camera (Conrady, 1919; Zhang, 2000). However, to meet the stringent requirements for precision in our predictive system, we improved upon existing procedures by implementing a novel feature detection algorithm leveraging the rotational symmetry of chessboard intersection points, and a method for finely estimating minute geometric imperfections during printing of our chessboard grids (see Appendix A). Together, these modifications of traditional calibration procedures yield a 96% reduction in calibration error, reducing the errors in single camera distortion correction from ~0.1 to ~0.02 pixels.

Along with image distortion correction, calibration of the precise relative position and orientation between a pair of stereo cameras is required for triangulating the 3D location of an object in the room given its 2D detected location in pixels on each camera's image. We used synchronously captured images of our chessboard grids in varied orientations and positions in the room to calibrate these stereo parameters; our implemented modifications of traditional calibration methods yield a 92% reduction in the stereo parameter calibration, reducing the errors from ~0.4 pixels across both cameras to ~0.1 pixels.

Given the large distance of 13-14 ft between the rim and the location of shot release, we tracked shot release with a pair of cameras separate from the pair used to track the ball's approach at the rim and shot outcome. After calibrating for the image distortion in each camera and stereo parameters for each pair of cameras, we did a final procedure to align the orientation of the hoop space camera measurements to the shooter space camera measurements. The procedure, outlined in Appendix A, fits a single rotation parameter and three translation parameters for the hoop space with an overall precision of ~0.250 mm.

1.3.2 Tracking ball motion

Calibration of image distortion, stereo camera geometry, and alignment parameters allows conversion of object positions detected in 2D pixel space on synchronous images in each camera to 3D world coordinates using simple projective geometry. The following subsections outline the procedures used for tracking the ball's position and velocity (linear motion), as well as the ball's rotational orientation and spin rates (rotational ball motion).

Tracking linear motion of the ball

State of the art camera tracking systems often employ the use of markers, such as small plastic spheres painted with a reflective paint, which need to be attached to the object or person being tracked. As this approach would likely disturb shooters, affect the flight of the ball, and would likely not be durable to physical collisions of the ball with the rim, backboard, and floor, we instead treat the ball itself as a large spherical marker and track its position in each image as its area centroid in the image. To facilitate this approach, we choose to use a dark ball against a bright white background which maximizes the contrast between the ball and the background. We then identify the pixels belonging to the ball as the darkest pixels in the image, selected by thresholding the image based at a specific brightness level. To improve robustness to stray detected pixels in distal parts of the image sometimes caused by flickering lights or pixel noise, we calculate the centroid in a few iterations, with each subsequent iteration windowing the image more tightly around the estimated ball centroid.

An illustration of ball centroid detection in each of the two shooter space cameras is shown as the cyan dots and traces in Figure 1.2 a & b, which are overlaid on synchronously captured views of the ball after shot release in each camera. After the ball is detected in each image, the detected 2D location in each image is triangulated using calibrated stereo parameters into a 3D position in world space. While we do not have a "gold standard" by which to compare our measurements, we estimate the precision of our 3D tracking of the ball's position by simply noting that physics requires the detected motion to be smooth. For 12,000+ shots, we estimated how smooth the tracked 3D positions were by fitting a straight line to 5

successive estimates of ball position sampled at 80 Hz, and quantifying the variance of the residuals. The distributions of estimated shot precision are shown in Figure 1.2c. Despite a ~2mm pixel size in images of the ball, the combined estimated precision in tracking the ball's linear position is ~60 µm on average.

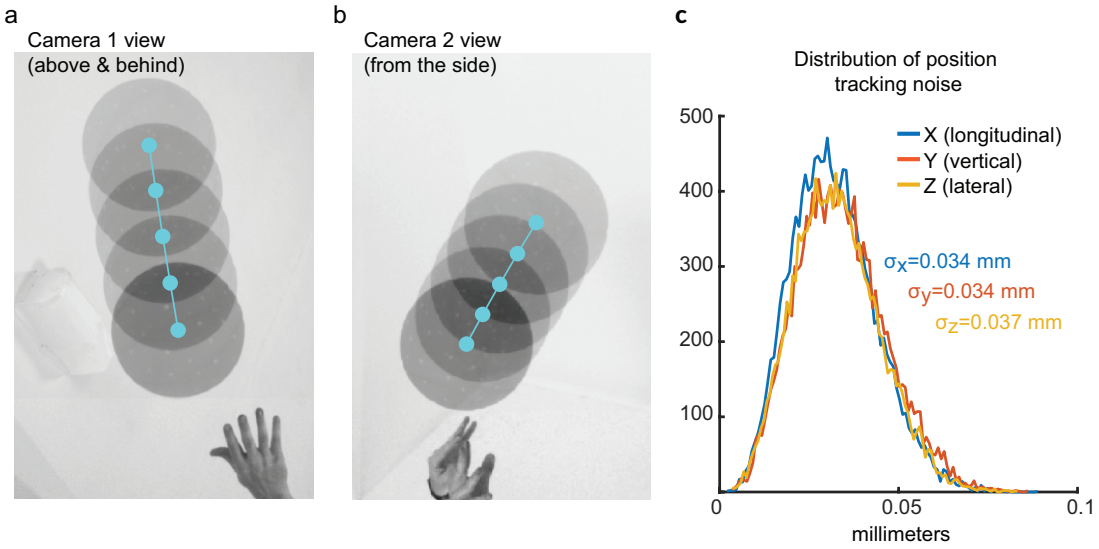

Figure 1.2: Tracking linear motion of the ball
(a-b): Illustration of ball centroid tracking in single camera images for 5 consecutive frames immediately following release of the ball from the shooters hand. Camera 1 views the ball from above the shooter looking down, while Camera 2 views the ball laterally from the side. Detected ball centroid for each frame is shown in blue. To aid visualization images of the ball from each frame have been overlaid on a single image, which shows the hand positions only for the first frame after detected release.
(c): Histograms of the estimated tracking noise in 3D after combining the 2D detected centroid positions illustrated in **a** and **b**. This precision tracking noise is a small fraction of the ~2mm pixel size.

Tracking rotational ball motion

Most experienced basketball shooters impart some spin, usually backspin, on the ball when shooting. From physics it is well known that rotation of ball while it moves through the air will create a lift force in a direction perpendicular to the current velocity vector of the ball. For backspin, this would cause a mostly upward lift force as the ball travels towards the hoop, though as the ball rises towards the peak of its trajectory the force will also push it back towards the shooter, and as the moves downward after the peak the lift force will push it somewhat forward. The exact direction, magnitude, and overall effect of the lift force depends both on the rate and direction of the spin (e.g. for pure backspin, spin around the

longitudinal and vertical axes would be zero). As shot-to-shot and shooter-to-shooter differences in spin characteristics are likely to affect the outcome of shots, we aimed to precisely track the spin characteristics of each shot.

To track the rotation motion of the ball, we track frame-by-frame changes in positions of dots on the surface of the ball relative to the ball's 3D centroid. To facilitate this, we painted 93 irregularly spaced ~1cm diameter gray dots on the surface of the dark ball. In particular, the irregular spacing of dots allows for each dot to be uniquely identified by its collection of distances to nearby dots. Figure 1.3a shows an image of the basketball with painted gray dots. The specific gray color used was chosen to appear in images as a brightness brighter than the dark ball, but considerably darker than the bright background, so as not to interfere with the brightness threshold used in processing images for overall ball centroid calculation. To detect the dots in single camera images, we first enhance the contrast specifically for small dots in the image by applying a 2D convolution to the image with a 10x10 pixel kernel defined to have a bright center and dark surround region (similar to the appearance of dots on the ball's surface). We then threshold this filtered image, which separates the dots from the background as clusters of detected pixels and apply a connected components analysis to separately get the centroid of each detected individual dot.

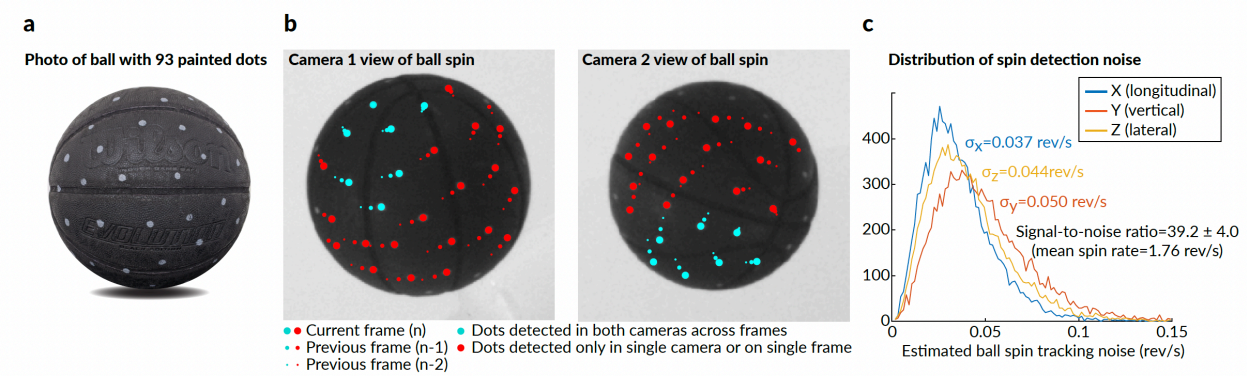

Figure 1.3: Tracking rotational ball motion using dots on the ball surface

(a): Dark basketball painted with 93 irregularly spaced gray dots whose brightness level in images is between in between the brightness of the ball and the background

(b): Illustration of 2D dot detection in individual cameras and matching across cameras for 3D detection. Images of the ball from camera 1 and camera 2 are overlaid with colored dots indicating the detected location of dots on the ball's surface. A camera-to-camera matching procedure is used to pair dots which appear in view of both cameras (blue dots), as opposed to those which are only visible in a single camera (red dots). Trailing dots of decreasing size indicate rotation of the ball by showing the history of dot locations relative to the ball's centroid.

(c): Histogram of noise in spin detection. Precision for detecting spin rate was ≤0.05 rev/s.

Once the dots are detected in each camera's individual images, converting their detected 2D positions in each image to 3D positions requires properly pairing the dots across the images from both cameras. We do this by a camera-to-camera matching process which evaluates all possible dot pairings across cameras with two criteria: (1) the combined reprojection error for the proposed pairing must be less than 0.05 pixels, and (2) the proposed pairing must produce a 3D position that is within 1mm of the ball's surface. Here, reprojection error is an error which reflects the degree of correspondence between information combined across two cameras. Specifically, it is the difference between the actual detected 2D location of a feature and the expected 2D location that feature if the estimated 3D position were reprojected back into the 2D space of the camera's image. Figure 1.3b shows an illustration of dot detection in single camera images, where the position of detected dots not matched across cameras are marked with red dots, and the position of dots matched across cameras are shown in blue. On average 12-14 dots which match across cameras are detected in each pair of captured images.

Using a library of preliminary shot data, we build a 3D map of dot locations on the ball using the detected 3D dot locations which meet the camera-to-camera matching procedure. Using this 3D map of n dots, we create a reference library of n x n pairwise distances between dots, where the collection of pairwise distances for each dot is treated as its uniquely identifiable signature. To estimate the rotational motion of the ball, we start by calculating the pairwise distance between all detected dots in 3D from each pair of images. We then compare these detected pairwise distances to reference library to uniquely identify each detected dot as a particular dot at a particular location on the ball's surface. We then calculate the rotation between the current detected dot positions and their position in the reference geometry, which gives us a measure of the rotational orientation of the ball on a given frame pair. Obtaining this rotational orientation on multiple frame pairs allows us to calculate the direction and magnitude of the ball's rotation rate in coordinates aligned with its linear velocity, which we generally refer to as the extrinsic spin rate. Distributions of extrinsic spin rate estimation noise are shown in Fig. 1.3d. Our overall precision in determining the ball's extrinsic spin rate is approximately 0.08 Hz. We can also calculate the direction of the ball's rotation relative to its geometry (i.e. is it rotating across the laces or along some other direction). We generally refer to the rotation in intrinsic coordinates of the ball's geometry, along with the initial rotational orientation of ball, as intrinsic spin information.

1.4 Classifying shot outcome

Aiming to train predictive models for shot outcome, we start by investigating the structure of shot outcome in a large dataset of tracked shots. Using a pair of synchronized stereo cameras with calibrated parameters for precise 3D tracking of the ball, we track the ball's trajectory as it approaches the rim for 50,000+ shots. While traditionally shot outcome has been defined simply by whether the ball goes through the hoop, we define a continuous measure of shot accuracy yields a much finer-grained estimate of how "good" a shot was. Further, we use this continuous measure of shot accuracy to develop a refined definition of binary shot outcome, which we call "direct makes".

Measuring shot endpoint and approach angle

As define a continuous measure of shot accuracy as the projected intersection point of the ball's trajectory with the horizontal plane at the rim height, if it were allowed to continue along its trajectory without rim or backboard contact (Fig 1.4a). We refer to this projected landing position of the ball as the "shot endpoint", and the landing position relative to the rim position as "shot endpoint error".

We estimate the ball's projected intersection point with the rim height horizontal plane using ball positions measured just before the ball contacts the rim or visually overlaps with the rim in the camera view. These tracked positions are projected forward in time using a simple equation for projectile motion, where the velocity is estimated as the average velocity over 3 frames of data (25 ms), and initial position is estimated as the average position over those same 3 frames. Because acceleration estimates from discrete position measurements are likely to be noisy for individual shots, we estimate acceleration as the average ball acceleration across shots. The average time horizon across which the trajectory is extended is ~100 milliseconds, and the estimated precision in determining the landing point is <0.5 mm. This imprecision is many times smaller than the average 10-20 cm variation in shot endpoints for individual shooters and is thus not a limiting factor in our analyses. Additionally, once we have the projected trajectory of the ball's approach towards the rim, we calculate the approach angle of the ball as the angle between its travel path and the horizontal plane just before it would reach the rim height.

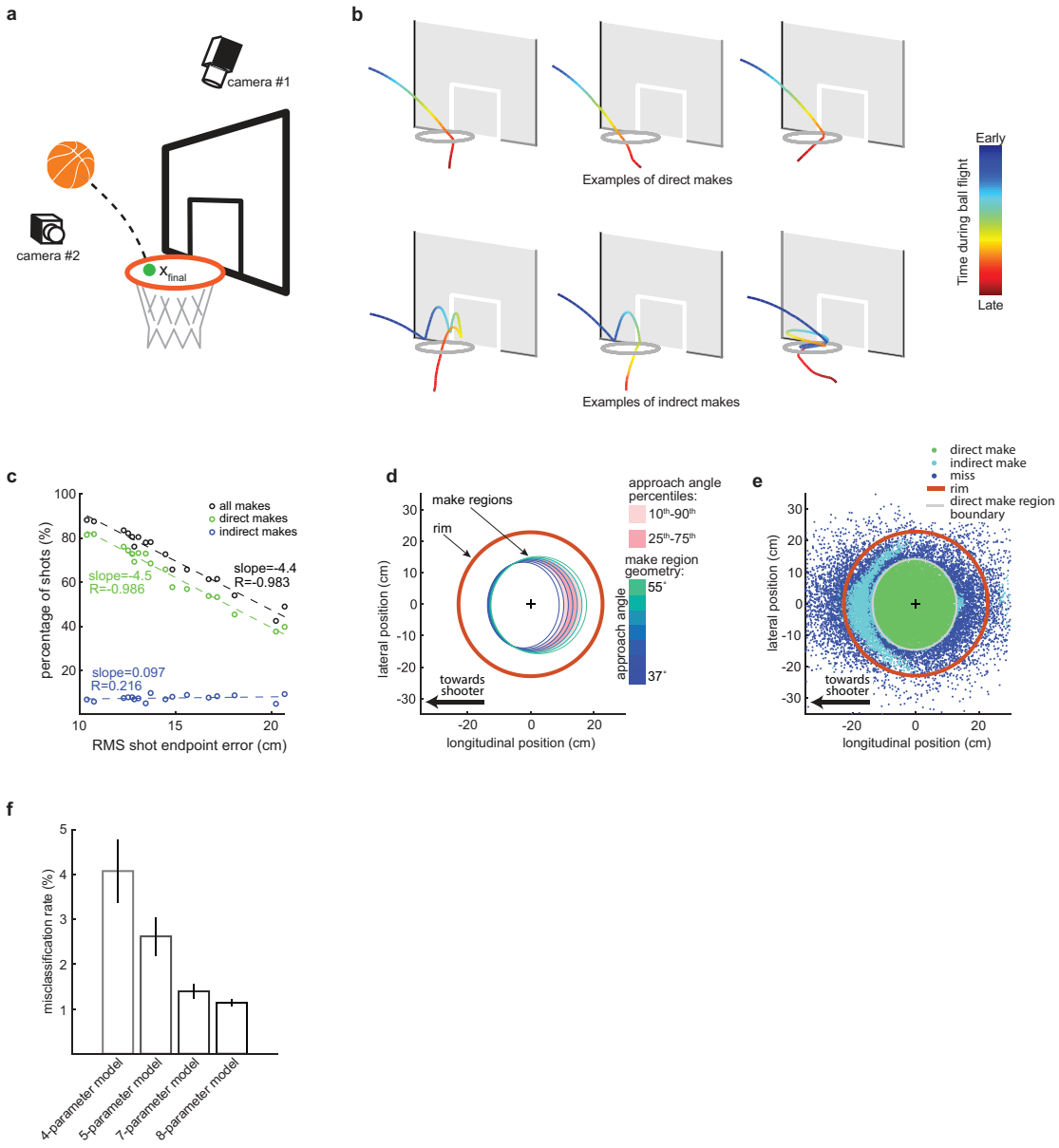

Figure 1.4: A system for precisely measuring and classifying shot outcome based on a decontaminated measure of shot success

(a): Illustration of shot endpoint tracking setup. Two synchronized cameras capture images of the ball as it approaches the rim at 80 frames/sec. Measurements of the ball's position immediately before it contacts the rim, or visually overlaps the rim in one of the camera views, are used to estimate shot endpoint, which we define as the projected point of intersection of the ball's trajectory with a horizontal plane at rim height. Images of the ball after it reaches the rim are used to determine whether the shot was a make or miss, and whether the ball passed quickly and directly through the rim, or whether it passed indirectly after bouncing upwards or around the rim.

Figure 1.4 (continued)

(b): Visualization of ball travel paths at the rim for direct and indirect makes. Example traces of tracked ball positions at the rim are shown in colored lines, with progressive change in trace color indicating the flow of time. The top row shows travel paths for direct makes, which are high accuracy shots that directly and quickly pass through the rim. If contact with the rim occurs at all, the post-contact ball trajectory is almost always downwards. In contrast, indirect makes (bottom row) are lower accuracy shots which fall through the rim after lucky bounces off the rim, with post-contact trajectories that are almost always upwards.

(c) Percent direct (green), indirect (blue) , and all (black) makes plotted against RMS shot endpoint error for 16 shooters with 3200+ shots each. Direct makes are a better indicator of shot accuracy ($R = -0.986$), compared to all makes ($r=-0.983$) due to removal of indirect makes, whose presence is unrelated to accuracy of the shot.

(d) Diagram demonstrating the modulation of direct make region geometry by approach angle of the ball. We define the direct make region with an elliptical boundary, based on shot endpoints and modulated by approach angle, to separate direct makes from misses and indirect makes. We include approach angle modulation of the make region geometry (blue to green) since steeper approach angles of the ball increase the effective size of the rim. The parameters for this classification model were estimated on a dataset of 50,000+ shots. As approach angles increase we see that the make region gets larger and moves closer to the back of the rim. Light and dark pink shading indicates the ranges for the middle 80% and 50% of make regions in the shot data used for fitting.

(e) Map of shot endpoints for misses (blue), indirect makes (light blue), and direct makes (green) in coordinates normalized to make region geometry. The map of shot endpoints relative to the rim (orange) and an average make region geometry (gray) shows clear separation of direct makes from other shots. Indirect makes are seen to occur in a mostly isolated low accuracy region of shot endpoints indicative of their limited ability to indicate shot accuracy.

(f) Cross-validated misclassification rate for different versions of make region classification model. Misclassification rates are shown for models using a make region with fixed geometry at rim height (4 parameters), fixed geometry with adjusted height (5 parameters), approach angle modulated geometry at rim height (7 parameters), and approach angle modulated geometry with adjusted height (8 parameters). To obtain cross-validated errors, models were fit on half of the available data and tested on the other half not used during fitting. Our system uses the 8-parameter classification model.

Direct makes as a refined binary shot outcome

We collected a dataset of 50,000+ shots from 16 shooters for which we tracked the outcome and endpoint on each shot. While looking at the tracked ball paths and the corresponding binary outcomes for several shots we identified two types of made shots, which we term direct and indirect makes. Direct makes are shots with high accuracy whose shot endpoints are close to the center of the rim, which either pass through the rim as a "swish" shot with no rim contact or pass through the rim quickly after brief contact with the rim. Indirect makes are usually shots with lower accuracy, whose shot endpoints errors are often similar to those of missed shots, but which still go through the rim by seemingly lucky bounces. These shots often bounce on the rim multiple times before passing through and take much longer to finally pass through the rim after contact compared to direct makes. Example tracked travel paths of actual shots falling into these two categories are visualized in Fig. 1.4b. Different from the direct makes, shown in the top row of Fig. 1.4b, the indirect makes bounce on the rim multiple times, or in the case of the third example, roll around the rim multiple times before falling through. We developed a criteria for identifying direct makes separate from indirect makes in the dataset. First, any shot which bounces upwards after contact with the rim is deemed an indirect make. Second, we measure the time it takes the ball to pass through the rim relative to when it arrives at the rim, and label any shot that takes longer than 150 ms to pass through the rim as indirect. Across the 16 shooters, with 3200+ shots each, we see a strong relationship ($r = -0.986$) between RMS shot endpoint error and percentage of direct makes, shown in green in Fig. 1.4c. In contrast, the relationship between shot endpoint error and percentage of indirect makes (shown in blue on Fig. 1.4c) is much weaker, with $R=0.37$, and is instead correlated positively with shot endpoint error, meaning that shooters with worse accuracy have more frequent indirect makes compared to better shooters. This suggests that direct makes are a more clarified indicator of shot accuracy compared to the traditional definition of a make, which includes direct and indirect makes.

A model for classifying direct makes

As indirect makes are not indicative of shot accuracy, we refine the definition of binary shot success to only count direct makes as success, and so we develop a classification model which uses an elliptical boundary to separate direct makes from both indirect makes and misses based on shot endpoint position. Since direct makes tend to be close to the center of the rim, the model assumes an elliptical region inside the rim, called the direct make region, with shots projected to land inside the region predicted as high-accuracy direct makes, and those projected to land outside the region predicted as lower-accuracy shots which could be misses or indirect makes. We fit the geometric parameters describing the make region, which includes 4 parameters corresponding to lateral centerpoint, longitudinal centerpoint, lateral width and longitudinal width, using built-in gradient descent functions in MATLAB and a custom loss function which is define below:

$$e_{direct\ make} = \begin{cases} 0, & when\ d \leq 1-\alpha \\ \dfrac{-d}{2\alpha} + 0.5, & when\ (1-\alpha) < d \leq (1+\alpha) \\ 1, & when\ d > (1+\alpha) \end{cases}$$

$$e_{miss} = e_{indirect\ make} = 1 - e_{direct\ make}$$

where $e_{direct\ make}$ is the error function for direct makes, and e_{miss} and $e_{indirect\ make}$ are the error functions for misses and indirect makes respectively; d is the normalized distance from the make region center, normalized by the radii of the ellipse, and α is the width of a region around the boundary inside which errors are a linear function of distance from the boundary. The continuous error region defined by width α facilitates using gradient descent methods to optimize the model parameters. Choosing an appropriate width of the continuous error region amounts to a tradeoff between sensitivity to individual shots and model generalization. When fitting the model parameters we use $\alpha = 0.1$.

As steeper approach angles of the ball make the rim appear effectively larger and more flat approach angles for "line-drive" shots make the rim appear effectively smaller, we also include parameters which

adjust the make region geometry based on the ball's approach angle. Specifically, we fit shifts to lateral and longitudinal centerpoint and gains to lateral and longitudinal make region width that were linear functions of approach angle, totaling 4 additional parameters. Finally, we include a parameter which adjusts the height of the horizontal intersection plane at shot endpoint is calculated, since the nominal 10ft rim height used for the plane may not afford maximum separation between the endpoints of direct makes and other shots. Using the optimized classification model, inspection of the make region geometries for a range of approach angles (blue to green ellipses in Fig. 1.4d) indeed shows that increasing approach angles cause expansion of the make region, but also shifts the make region further towards the back of the rim. We can use this classification model to better visualize the structure of shot outcomes by mapping shot endpoints for misses, indirect makes, and makes in coordinates normalized to the direct make region (Fig. 1.4e). Here we see that direct makes (green) are cleanly separated from other shots by the make region boundary (gray). Additionally, indirect makes occur in a mostly isolated low-accuracy region surrounded by misses, and for a vast majority of indirect makes, the plot indicates that a similar shot with slightly improved accuracy would likely be a miss. This result is somewhat counter-intuitive from the shooter's perspective, as indirect makes can outwardly appear to be shots with accuracies in between misses and direct makes. In fact, our analysis shows that shot accuracy is often worse for indirect makes than misses occurring close to the direct make region. Though counter-intuitive for shooters, this is consistent with the results in Fig. 1.4c which indicates that indirect makes are a poor proxy for shot accuracy.

Fig 1.4f summarizes the overall model performance for different versions of the classification model. While a 4-parameter model with a fixed make region geometry performs moderately well with a misclassification rate of 3.8%, inclusion of height adjustment and approach angle modulations reduces the misclassification rate to 1.2%, corresponding to ~68% improvement in classification.

Taken together, the elliptical classification model for a refined definition of shot success, which separates direct makes from other shots based on measured shot accuracy, allows the system to convert shot

endpoints into binary shot outcomes with high precision. To use this model during real-time experimentation, we design prediction algorithms focused on accurately predicting shot endpoint.

1.5 Predicting shot endpoint

Predicting whether a shot will be a direct make requires us to predict what the endpoint of the shot will be. After the ball is released from the shooter's hands, its trajectory is completely determined by the physics of projectile motion. Unlike body kinematics, which are idiosyncratic across shooters and have shown limited success in predicting shot outcome (Ammar et al., 2016; Malone et al., 2002; Mullineaux & Uhl, 2010; Özkan et al., 2019), projectile physics is consistent across all shooters and has been heavily studied and characterized. Thus, we focus on using early information about the ball's motion immediately after the release to predict where the ball will land, allowing us to use the elliptical classification model discussed in the previous chapter for to convert these predictions into binary shot outcomes.

1.5.1 Measuring the ball's initial state after release

To track the ball's state after release we use a pair of synchronized stereo cameras aimed at the region where shooters release the ball, which we carefully calibrated to allow precise 3D measurements of the ball's linear motion and rotational motion in the manner outlined in Chapter 1.3.2. Camera 1 was placed above and behind the shooter with a downwards view of shot release from a height of ~18 ft (tilted so that the view was approximately perpendicular to the motion of the ball). Camera 2 was placed on a lateral with a side view of ball release from a distance of 10 ft to the right of the shooter. Both cameras captured synchronized images with a resolution of ~2 mm/pixel at 80 frame/sec.

We estimate the initial state of ball for each shot by first tracking the ball's position (3 parameters) and rotational orientation (3 parameters) on each of two successive frames following separation of the ball from the shooter's hand. As the cameras capture images at 80 frames/sec, the time difference between frames is 12.5ms, and we use our estimates of linear and rotational position across the two frames to calculate estimates of the initial linear velocity (3 parameters), rate of rotation in world coordinates (3

parameters), and rotation direction in intrinsic coordinates of the ball's geometry (3 parameters) across the two frames.

Together, these 15 parameters describe the initial state of the ball after release. To fit a predictive model we combine these measurements with estimates of the shot outcome, including shot endpoint as described in the previous section, approach angle, and the categorical label for each shot indicating whether it was a direct make, indirect make, or miss. While we could attempt to fit a model directly using the parameters and shot outcomes as they are, we re-organize the data to simplify the prediction. We first align each shot and its corresponding endpoint into a shot-oriented coordinate system where the coordinates for each shot are rotated about the vertical (y) axis such that the initial lateral (z) velocity is zero. This yields a shot travelling only in a longitudinal (x) direction defined in the initial ball heading direction, and we applied this transformation to the external spin rate of the ball as well. Second, instead of predicting the actual lateral and longitudinal endpoint of each shot, we predict lateral and longitudinal excursions of the ball, which are defined as the distances between the initial position of the ball at release and the shot endpoint. This allows us to drop initial x- and z-position from the parameters describing the ball's initial state. Together, this yields 12 parameters describing the ball initial state, which are initial y-position, initial velocity (x & y), external spin rate (x, y, & z), intrinsic spin direction (x, y, & z), and initial rotational orientation (x, y, & z). These 12 parameters will be used in a model to predict the lateral and longitudinal ball excursions for each shot, and the approach angle for each shot which is necessary for predicting whether a shot will be a direct make. We use a dataset of 50,000+ shots from 16 shooters, with 3200+ shots from each shooter, to fit and evaluate models for predicting shot outcome.

1.5.2 A flexible regression model based on polynomial expansion of initial ball state

As our predictive model we decide to use a linear combination of terms from a polynomial expansion of the 12 parameters describing the ball's initial state to predict approach angle, lateral ball excursion, and longitudinal ball excursion. This gives our model large flexibility for representing various non-linear relationships between the response variables and the measured initial state. Polynomial functions would

also be well suited to describing the ball's physical motion since they are smooth functions, and we know that the physics of projectile motion predicts smooth motions of the ball. While we could use an actual physics-based model, a polynomial expansion of the initial state is likely to already include adequate representation of physics model terms. Further, the polynomial model could also allow fitting of nuanced relationships specific to basketball shooting and our setup. Examples of such relationships would be dependence of drag effects or the lift force on the orientation of ball laces, or slight misalignments between the hoop and shooter camera spaces. Here we use a linear regression on polynomial expanded terms to predict ball excursions and approach angle.

While polynomial expansion allows great flexibility in the predictive model by generating a host of possible relationships between the ball's initial state and the response variables. For a 4^{th} order polynomial expansion, the 12 parameters representing the ball's initial state are expanded into 1,820 polynomial parameters. However, many of these parameters are not likely to be meaningful for prediction, and inclusion of these extraneous parameters will impede precise estimation of the model coefficients, reducing model generalizability and exacerbating overfitting. Thus, use of the expanded polynomial model requires a parameter selection procedure to remove those which do not meaningfully predict changes in the response variables. We implement a parameter ordering process that sequentially orders parameters using their estimated with response variables while managing the widespread covariance among the polynomial. First, we sort all the polynomial parameters by their correlations with the response variable. The parameter with the highest correlation is selected first and removed from the list of remaining parameters. Critically, before selecting the next parameter we update the remaining parameters and the response variable by removing their linear relationship with the selected parameter using regression. The correlations for the updated parameters and updated response variable are then calculated again, and the process is repeated until all parameters have been sequentially ordered in this way.

In general parameters selected further along the ordering procedure are less related to the response variables. However, it is unclear based on the ordering process alone at number of parameters is optimal. To determine how many parameters to use in the predictive model we need to know how well predictions using some chosen number of parameters generalizes, and so we perform leave-one-out cross-validation (LOO) across participants in our dataset for different model sizes (Fig. 1.5a). Here we regress the response variable (i.e. longitudinal excursion), onto *n* ordered parameters for 15 subjects and test the performance of the fitted coefficients on the 16^{th} subject's data. We do this when leaving out each of the 16 subjects one at a time for a range of model sizes defined by the number of included parameters *n*. Fig. 1.5a shows the mean LOO errors across subjects for different model sizes. Initially as the number of parameters is increased the LOO errors rapidly drop. However, after a certain point the errors start to increase, which indicates that the model is overfitting the training data and not generalizing well to participants not included in the fit. The range of model size yielding the region of smallest cross-validation errors is used as an estimate of the number of parameters we can meaningfully fit given the size of our dataset.

The traces in Fig 1.5b indicate that approximately 100, 75, and 40 parameters should be used respectively for predicting longitudinal excursions, lateral excursions, and approach angle, and the corresponding across participant mean LOO prediction errors (±sem) are 36.2 ± 8.6 mm, 23.5 ±2.1 mm, and 0.75 ± 0.15 degrees respectively. Of the 16 subjects, one subject had abnormally large prediction error for all response variables (>5 IQR's from the median across subjects). When this subject is removed, the estimated prediction errors are 28.6 ± 4.3 mm, 23.8 ± 2.2 mm, and 0.62 ± 0.06 degrees respectively.

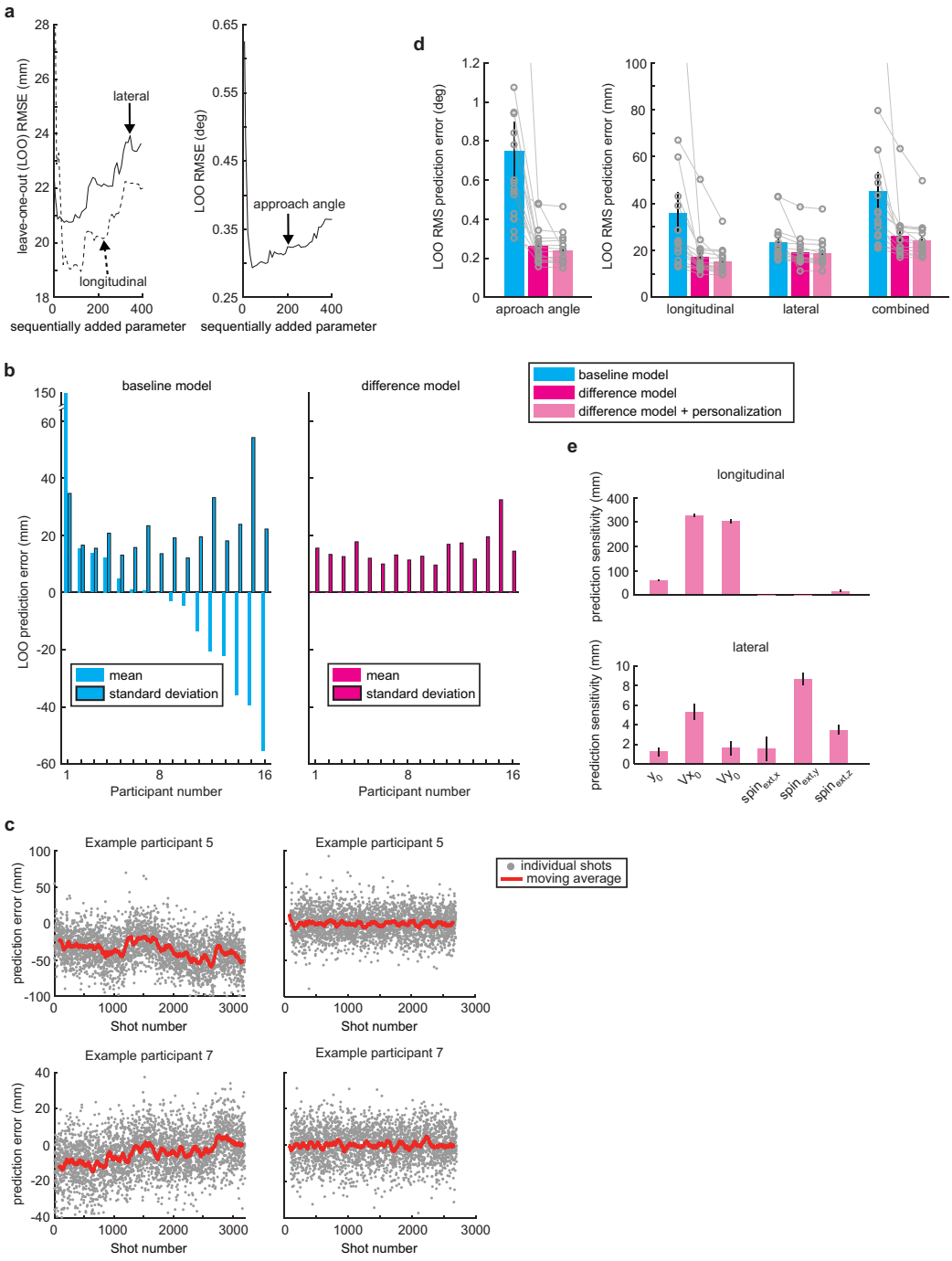

Figure 1.5: Prediction model for shot endpoint and approach angle based on initial ball state

We designed a predictive model which predicts a shot's landing position at the rim using parameters drawn from a 4th order polynomial expansion of a 12-dimensional representation of the ball's initial motion state. This representation of the ball's initial motion after it leaves the shooter's hands includes information about its initial position, initial linear velocity, rotational motion in coordinates of shot direction, rotational direction along the ball's intrinsic geometric coordinates, and initial rotational orientation of the ball.

Figure 1.5 (continued)

(a) LOO cross-validation by participant on models of increasing complexity. To determine model structure we performed LOO cross-validation on models which include progressively larger numbers of parameters. Parameters were added to the model in sequential order as determined by correlation-based method for ordering parameters based on their estimated importance to prediction of the response variables, which are lateral and longitudinal ball excursion (left panel) and approach angle of the ball at the rim (right panel). We chose the total number of parameters to use by finding the number of parameters in LOO cross-validation that produced the smallest prediction error.

(b) Dissection of longitudinal prediction errors into offset and variance reveals large inter-individual difference in prediction quality. Model prediction errors were dissected into their mean and variance components to investigate performance consistency across individuals. The base model (blue), which performs a simple regression on the selected parameters, produces large mean offsets (black outline) in prediction that are consistent within individuals and often similar in size or larger than the standard deviation (no outline). To combat this, we implemented a difference model (pink) which predicts the shot to shot changes separately from participant offsets. The difference model dramatically eliminates the individual mean offsets in prediction and also improves the variance of prediction errors, making prediction performance more uniform across shooters.

(c) Comparing base model and difference model performance in plots of prediction errors for example subjects. Base model (left) and difference model (right) prediction errors for two example subjects are shown in detail. Base model performance displays relatively large mean offsets which also appear to evolve over time. The difference model employs a moving average window of previous shots to compensate these time-evolving offsets which nearly eliminates effect of these offsets.

(d) Summary of prediction improvements from difference model. For approach angle, longitudinal excursion, and lateral excursions the difference model (dark pink) vastly improves performance for all individual participants (gray) over the base model (blue). We include a modification of the difference model (light pink) which incorporates participants' baseline shooting data into the model fit to personalize the model, which mostly benefits longitudinal predictions.

1.5.3 Improving model performance with separate components for offset and variance prediction

When evaluating the performance of the predictive models we chose to use LOO prediction for participants, instead of other variations of cross validation, because it better reflects the way in which we would implement the model in practice. Particularly, large differences in predictive performance across individuals would introduce noise in experimentation as inconsistency in feedback quality across participants. We looked at how well the prediction was working across individuals by calculating the mean and standard deviation of the residuals from our prediction. Fig. 1.4c shows the mean and standard deviation of prediction errors for longitudinal ball excursion in individual subjects, sorted by the mean prediction error. Both the mean and standard deviation of predictions varies significantly across individuals. Indicative of the size the differences, there is a factor of 4 difference in standard deviation of prediction errors across individuals with the lowest and highest prediction error variance. Unexpectedly, the mean prediction errors are also quite large for several individuals, often equal or larger in magnitude compared to the standard deviation. Further, they are quite different across individuals which we did not *a priori* expect since the physics of ball motion is the same across participants. Upon closer inspection, we saw that not only were the participant mean prediction offsets consistently different for subjects, they also displayed some time-dependent drifts throughout the set of shots within individuals. Fig. 1.4c shows, for two example subjects (P5 and P7 from Fig. 1.4b), how the mean offsets (red) evolve throughout the shooters' 3200 shots.

The differences seen in model performance across individuals in both the mean of prediction errors and their variances indicated that the current model was likely ill-equipped to fit both shot-to-shot changes and the time-evolving participant specific offsets in the data. Thus, we implemented a version of the polynomial regression model which separates prediction of shot-to-shot changes and the time-evolving offsets. We call this model the "difference model" and now refer to the original model described above as the "base model". In the difference model, we obtain the coefficients for predicting the shot-to-shot changes by regressing the shot-to-shot changes in response variables onto the corresponding shot to shot

changes in the polynomial parameters. As these predictions carry no information about the mean prediction level, we estimate the mean prediction offset using a moving average window of *w* previous shots.

We first determined structure of the difference model by calculating the ordering of the parameters, which are now pre-processed as the shot-to-shot change instead of their nominal values, and performed LOO cross-validation across participants in a manner analogous to what was done for the base model. Additionally, we calculated the average RMS prediction error across subjects in the LOO analysis for a range of window sizes for the moving window offset compensation component of the model and determined that a window size of approximately 100 previous shots is optimal. We compare the performance of the difference model to the base model by once again looking at the mean and standard deviation of predictions for individual subjects (Fig 1.4d). The mean offsets for individuals are massively reduced in magnitude when using the difference model, so much so that they are hardly visible on the plot when the y-axis limits are the same as those used for the base model in Fig. 1.4b. Where as the average magnitude of the mean offsets in prediction for shooters was 24 mm in the base model, all mean offsets for individual shooters in the difference model are below 0.3 mm. Inspection of the prediction errors on individual shots within shooters shows that the moving window offset compensation adequately reduced the time-evolving drifts in mean prediction errors within shooters. For participants 5 and 7 we can compare LOO prediction errors coming from the base model and difference model across panels c and e in Fig. 1.4, which illustrates the reduction in contribution of the time-evolving offsets. Combining LOO prediction errors across participants (51,000 shots), the overall improvements when switching from the base model to the difference model expressed as partial R^2 are 0.85, 0.33, and 0.91 for longitudinal excursion, lateral excursion, and approach angle predictions. In particular, the difference model dramatically reduces the errors for multiple shooters, including the shooter with very high base model prediction errors for longitudinal excursion.

The difference model greatly improves the prediction performance of the system. Fig 1.4f summarizes the improvements in approach angle and ball excursion prediction. We include in the summary an improved version of the difference model where the first 500 shots from each shooter are included in the training data when performing LOO cross-validation. This version of the model would be a realistic implementation in experiments where shooters take a baseline set of 500 shots before any predictive feedback is given to them, allowing us to "personalize" the model to their shot style. For the difference model with personalization, we report the LOO prediction errors only considering the 2700 shots not included in the training dataset. Improvement from personalization is small for predicting lateral excursions (partial-R^2=0.06), and moderate for predicting longitudinal excursions (partial-R^2=0.29) and approach angle (partial-R^2=0.21). The average RMS LOO prediction errors for the difference model with personalization (mean ± sem) are 15.0 ± 1.4 mm for longitudinal excursions (R^2=0.994), 19.0 ± 1.54 mm for lateral excursions (R^2=0.45), and 0.24 ± 0.02 degrees for approach angle (R^2=0.997).

While the R^2 for longitudinal and approach angle predictions are high indicating good predictive performance, it is unclear whether the moderate performance for lateral predictions is from model inadequacy or because there is less predictable variance. Assuming that noise in measuring the ball's position on each frame is independent, our estimated of 40 μm precision in ball position would yield approximately 4.5 mm lateral prediction errors, which would account for 6% of the variance in our prediction errors from the difference model.

We use the difference model to understand which release parameters affect shot outcome the most by calculating the sensitivity of excursion predictions to changes in the release parameters. Sensitivity was calculated as the change in predicted excursion size across a range of two standard deviations for each release parameter, where the standard deviation for each parameter was estimated across subjects (Fig. 1.4g). For longitudinal excursions (top plot) we see that initial x- and y-velocities vastly dominate how far the ball will travel. This is consistent with previous work (Tran & Silverberg, 2008) based on the simulated physics of ball flight. However, it is interesting to note that the variations in z-spin rate, which

is the rate of backspin, has a sensitivity of ~17 mm. Compared to the sensitivities for Vx0 and Vy0, which are 329 mm and 304 mm respectively, this seems rather small. However, this sensitivity is 2x larger than any sensitivity for lateral excursions. For lateral predictions, the sensitivities are smaller than those for longitudinal, consistent with a more moderate R^2. Interestingly, lateral predictions are most sensitive to spin around the vertical (y) axis (frisbee-style spin). This makes sense since spin around the vertical axis as the ball moves forward would produce a lateral lift force pushing the ball off course to the left or right.

1.5.4 Role of spin in basketball shot mechanics

Spin imparted on the ball governs unpredictability in shots

Looking at the model performance for individual shooters (Fig. 1.4 b & d) we see a drastic improvement in prediction error offsets when using the difference model, and a clear improvement in the variance as well. While the difference model predictions are more consistent across shooters than the base model, there are still consistent differences across shooters in the variance of prediction errors. For example, in Fig. 1.4d we see that the standard deviation of longitudinal prediction errors for P15 is 32.5 mm, but <10 mm for shooters P6 and P10. To understand what could be driving such differences in prediction quality we looked at relationships between RMS prediction error and parameters describing the ball's initial state. Strikingly, as shown in Fig. 1.5a, we found that the participants' average total spin rate (combined across x, y, and z directions) was correlated with the average magnitude of prediction errors for both lateral and longitudinal predictions (r = -0.84 for longitudinal excursion prediction, r = -0.71 for lateral excursion prediction). This would suggest that the predictability of shots might be modulated by the amount of spin imparted on the ball. To see this relationship at a finer level of detail, we combined shots across participants and binned them by their total spin rate into 50 evenly spaced bins, and calculated the average magnitude of prediction errors for longitudinal and lateral excursions (Fig. 1.5b). Here we can see a cleaner relationship between predictability and total spin rate, with correlations of r = -0.92 and r = -0.82 for longitudinal and lateral excursions respectively.

This relationship between shot predictability and amount of spin suggests that higher spin rates may be reducing some inherent unpredictability in the shot's landing point. The plot in Fig. 1.5b indicates that this effect may reach an asymptotic unpredictability of ~15mm in the lateral direction and possibly 7-8 mm in the longitudinal direction. A likely physical mechanism explaining this relationship is the phenomenon of "knuckling", which gets its name from the infamous "knuckleball" pitch in baseball. In this pitch, a pitcher grips the ball with the knuckles of their fingers instead of their fingertips, critically reducing the amount of spin imparted on the ball which normally stabilizes the airflow around the ball. As this low-spin pitch travels through the air it tends to irregularly shift laterally and vertically due to unstable aerodynamic forces around the ball, which are normally partially stabilized by spin, making it hard for the batter to predict where the ball will cross home plate (Texier et al., 2016; Watts & Sawyer, 1975). In basketball, this phenomenon would lead an increase in random, unpredictable motions of the ball during flight for shots with lower spin. Knuckling is reported to occur for situations where the Reynold's number ranges between $3x10^4$ and $6x10^5$, where Reynolds number is a non-dimensional number describing flow characteristic around or through an object and is calculated as

$$Re = \frac{\rho V D}{\mu}$$

where ρ is the density of the fluid, μ is the fluid viscosity, V is the flow velocity, and D is the characteristic length. In our dataset, the middle 95% of shots have velocities ranging from 6.25 to 7 m/s, which correspond to Reynolds numbers of $1.02x10^5$ and $1.14x10^5$. This places basketball shooting well within the range of Reynolds numbers for knuckling phenomenon, and so it is likely that the increased predictability of high spin shots is driven by more stable airflow and less random motion from knuckling. As individual shooters tend to differ in their mean spin rate, this phenomenon is most likely the driving factor behind individual differences in shot prediction quality.

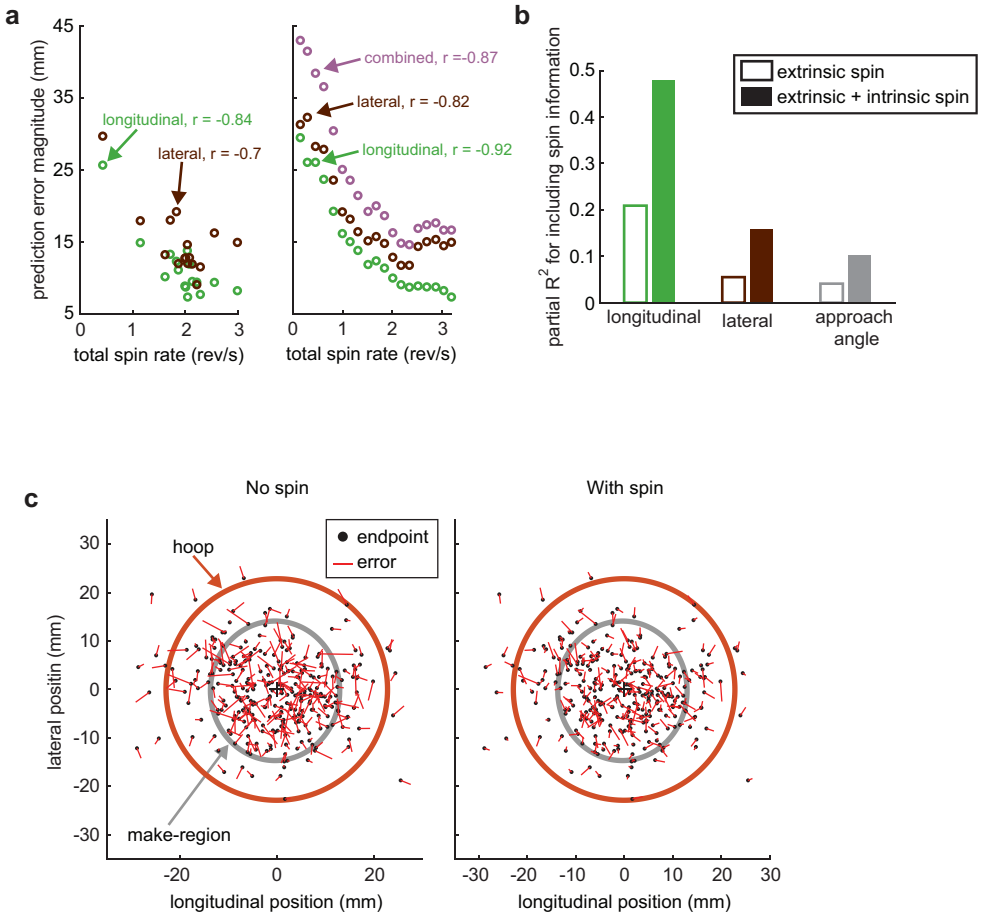

Figure 1.6: Influence of spin on prediction of shot endpoints

Motivated by consistent differences prediction quality across shooters, we investigated whether differences in prediction quality were related to any of the measured parameters describing initial ball motion state. We found that predictive performance increases with total spin rate, likely caused by a "knuckling" effect where rapid changes in airflow around the ball induce lateral forces which randomly push the ball off course during travel through the air. **(a)** Average prediction quality for individual shooters depends on total spin rate. For longitudinal (green) and lateral (black) prediction errors, we find a relationship between the magnitude of prediction errors and average participant total spin rate (left panel), which is the linear combination of extrinsic spin rate for x, y, and z directions. This suggests a stabilizing effect of ball spin on the ball's flight path. To see the relationship between total spin rate and prediction quality more cleanly, we pooled data across participants and binned errors by total spin rate. While both longitudinal and lateral prediction quality is affected by total spin rate, calculated correlations suggest there may be a stronger stabilizing effect in longitudinal errors affect. Bins containing <150 shots are not shown or included in analysis, and correlations shown are Pearson's correlation coefficient.

Figure 1.6 (continued)

(b) Improvement in predictive model performance when including intrinsic and extrinsic spin information. To better understand the role of spin in prediction, we quantified model improvement when including intrinsic and extrinsic spin, which are expressed on the plot as partial R^2 values. Inclusion of spin information is most beneficial to longitudinal predictions, which makes sense since most ball spin is mostly backspin which causes lift forces that effect the distance the ball travels in the longitudinal direction.

(c) Improvements in prediction error from inclusion of spin information leads to fewer misclassifications. Measured shot endpoints (black) for a random subsample of shots are plotted with stems showing their prediction errors (red) when either no spin information was included when fitting the predictive model (right), or when extrinsic and intrinsic spin information were included (left). Stems which cross the make region boundary (gray) amount to classification errors. The smaller prediction errors when spin information is included (right) contributes to fewer misclassifications in predicting shot outcome.

Inclusion of extrinsic and intrinsic spin information improves prediction performance

Our system uses measurements of the ball's rotational orientation relative to an arbitrary reference orientation on each frame and generates estimates of extrinsic and intrinsic spin information. Extrinsic spin parameters include the ball's rotation vector in coordinates defined by the shot heading direction. Intrinsic spin parameters include the ball's rotation direction in reference to ball's geometry, and the initial rotational orientation of the ball at the moment of release. We investigated the contributions of these two types of spin information by first fitting a model with no spin information, and then calculating how much of the remaining variance is explained when including extrinsic spin information only, or extrinsic and intrinsic information (see Fig. 1.5c). The extrinsic spin information is expected to be most useful in determining the effect of lift forces, which are forces perpendicular to the linear velocity of the ball arising from spin. When extrinsic information is included in the model, we see small improvements for predicting lateral excursions and approach angles corresponding to partial-R^2 values of 0.04 and 0.05, respectively. However we see a reduction of approximately 20% for longitudinal excursion prediction. Including both intrinsic and extrinsic spin information again yields a much larger improvement for longitudinal predictions, where the we a 50% decrease in the remaining variance compared to a model with no spin. Improvements for lateral excursions and approach angle were again small, with partial R^2 of

approximately 0.14 for both. The larger improvements seen for longitudinal excursions is expected because the predominant component of ball spin among shooters is backspin, which affects longitudinal motion of the ball.

The improvement in model prediction from inclusion of spin information greatly improves our system's ability to predict whether shots will be direct makes or not. Fig. 1.5 d & e demonstrate this improvement in coordinates relevant to classification. The dots indicate shot endpoint locations for 300 shots, while the stems for each dot indicate the size of prediction errors when using a model without spin information (d) or a model with extrinsic and intrinsic spin information (e). Lines crossing the average make region boundary shown in gray are likely to be misclassified, and so the increased error sizes in the no-spin model lead to higher numbers of misclassifications.

1.5.5 Summary

In this section we described the development of a predictive model for longitudinal shot endpoint, lateral shot endpoint, and approach angle based on a polynomial expansion of 12 parameters describing the ball's initial state immediately after release. The model, which we call a "difference model", implements separate components for predicting the shot-to-shot changes and the time-evolving offsets in predictions for individual shooters. The overall accuracy in for our model is 1.5 cm for longitudinal shot endpoint (R^2=0.994), 1.9 cm for lateral shot endpoint (R^2=0.45), and 0.23 degrees for approach angle (R^2=0.997).

1.6 References

Ammar, A., Chtourou, H., Abdelkarim, O., Parish, A., & Hoekelmann, A. (2016). Free throw shot in basketball: kinematic analysis of scored and missed shots during the learning process. *Sport Sciences for Health*, *12*(1), 27–33. https://doi.org/10.1007/s11332-015-0250-0

Brudner, S. N., Kethidi, N., Graeupner, D., Ivry, R. B., Taylor, J. A., & Taylor, J. A. (2016). Delayed feedback during sensorimotor learning selectively disrupts adaptation but not strategy use. *J Neurophysiol*, *115*, 1499–1511. https://doi.org/10.1152/jn.00066.2015.-In

Conrady, A. E. (1919). Decentered Lens-Systems. *Monthly Notices of the Royal Astronomical Society*, *79*(5), 384–390.

Malone, L. A., Gervais, P. L., & Steadward, R. D. (2002). Shooting mechanics related to player classification and free throw success in wheelchair basketball. *Journal of Rehabilitation Research and Development*, *39*(6), 701–709.

Mullineaux, D. R., & Uhl, T. L. (2010). Coordination-variability and kinematics of misses versus swishes of basketball free throws. *Journal of Sports Sciences*, *28*(9), 1017–1024. https://doi.org/10.1080/02640414.2010.487872

Nakano, N., Inaba, Y., Fukashiro, S., & Yoshioka, S. (2020). Basketball players minimize the effect of motor noise by using near-minimum release speed in free-throw shooting. *Human Movement Science*, *70*(October 2019), 102583. https://doi.org/10.1016/j.humov.2020.102583

Özkan, D. G., Pezzetta, R., Moreau, Q., Abreu, A. M., & Aglioti, S. M. (2019). Predicting the fate of basketball throws: an EEG study on expert action prediction in wheelchair basketball players. *Experimental Brain Research*, *237*(12), 3363–3373. https://doi.org/10.1007/s00221-019-05677-x

Roach, N. T., Venkadesan, M., Rainbow, M. J., & Lieberman, D. E. (2013). Elastic energy storage in the shoulder and the evolution of high-speed throwing in Homo. *Nature*, *498*(7455), 483–486. https://doi.org/10.1038/nature12267

Schween, R., & Hegele, M. (2017). Feedback delay attenuates implicit but facilitates explicit adjustments to a visuomotor rotation. *Neurobiology of Learning and Memory*, *140*, 124–133. https://doi.org/10.1016/j.nlm.2017.02.015

Silverberg, L., Tran, C., & Adcock, K. (2003). Numerical Analysis of the Basketball Shot. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, *125*(4), 531–540. https://doi.org/10.1115/1.1636193

Texier, B. D., Cohen, C., Quéré, D., & Clanet, C. (2016). Physics of knuckleballs. *New Journal of Physics*, *18*(7). https://doi.org/10.1088/1367-2630/18/7/073027

Tippetts, B., Lee, D. J., Lillywhite, K., & Archibald, J. (2016). Review of stereo vision algorithms and their suitability for resource-limited systems. *Journal of Real-Time Image Processing*, *11*(1), 5–25. https://doi.org/10.1007/s11554-012-0313-2

Tran, C. M., & Silverberg, L. M. (2008). Optimal release conditions for the free throw in men's basketball. *Journal of Sports Sciences*, *26*(11), 1147–1155. https://doi.org/10.1080/02640410802004948

Watts, R. G., & Sawyer, E. (1975). (1975) Aerodynamics of a knuckleball.pdf. *American Journal of Physics*, *43*(11), 960–963.

Young, R. W. (2003). Evolution of the human hand: the role of throwing and clubbing. In *J. Anat* (Vol. 202).

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), 1330–1334. https://doi.org/10.1109/34.888718

Chapter 2: Predictive delay cancellation in basketball free-throw shooting

2.1 Abstract

A growing body of work suggests that delays between a movement and its consequent feedback impairs motor learning and motor performance. While most existing work focuses on latencies of hundreds of milliseconds, little is known about how the motor system learns across longer delays on the order of seconds, such as those ubiquitous in skilled throwing motions in sports like football and baseball. Free-throw shooting in basketball is well-positioned to be a model system to study such latencies in complex movements due to its widespread popularity, its unique existence in sports as measure of skill unfettered by defensive pressure, and its consistent task specifications across all games and shooters. Here we trained shooters in a system capable of circumventing the inherent ~1 s feedback delay in free-throw shooting by providing shooters with highly accurate predictive feedback about shot outcome just 35 ms after they release the ball. Across six days of training comprising 3,200 shots, we did not detect a significant change in overall performance in shooters trained on time-advanced feedback relative to controls. However, we did see a small improvement in the longitudinal alignment of shot endpoints with the success region specific to shooters receiving time-advanced feedback, suggesting the potential of time-advanced feedback to improve shooting performance in refined experimental designs.

2.2 Introduction

Many movements we perform every day are accompanied by immediate temporal feedback about the outcome relative to their motor goals: for example, when reaching for your coffee mug across the table, feedback about whether you successfully grabbed the handle is available to you nearly instantly as you grasp the handle. However, many other movements, including ones that require use of human computer interface (HCI) systems like computer mice and video game controllers, incorporate small delays between a person's movement and the corresponding outcomes due to processing time.

Though these feedback delays are usually only tens to hundreds of milliseconds, a growing body of work suggests that they can impact motor performance across a variety of tasks. For example, latency-induced impairments in performance have been observed in expert-driven robotic surgeries (Kumcu et al., 2017), classical conditioning tasks in animals that involve pairing of airpuff-evoked blinks with tones (Kehoe et al., 2009; Smith et al., 1969), as well as arm reaching movement tasks in which humans learn to alter their hand paths to compensate for an experimentally altered relationship between hand motion and visual feedback of that hand motion (Brudner et al., 2016; Schween & Hegele, 2017). Such decrements in motor performance are consistent with a recent neurological examining Purkinje cell plasticity in cerebellar cortex, which is thought to play an important role in error-feedback driven motor learning (Najafi & Medina, 2013; Ohmae & Medina, 2015). Plasticity in these cells, driven by precise coincident timing between activity in climbing fibers carrying error information and parallel fibers carrying sensory information, was found to be sensitive to latencies as small as 20 ms (Suvrathan et al., 2016).

A key insight gleaned from the studies outlined above is that they examined motor performance in response to latencies that were restricted to a range of tens to hundreds of milliseconds. Thus little is known about how longer delays affect motor learning, such as those that occur naturally in ball sports and skilled throwing tasks. While throwing a football, for example, the inevitable delay between completion

of the throwing motion and target-contact is on the order of seconds, which is significantly longer than latencies typically found in HCI systems or those used in previous work.

Here we sought to determine the effects of long latencies naturally present in a skilled throwing task. Such tasks are ubiquitous across a range of sports and activities such as frisbee, baseball or basketball, and have even been hypothesized as being a driver for human evolution (Roach et al., 2013; Young, 2003). Therefore, although such skills are complex and are often honed over years, examination of them offers key insight into the formation of a fundamental movement widely used during environmental interaction.

In particular, we examined free-throw shooting in basketball, which contains a natural delay of 900-1000 ms between completion of the shooting motion and arrival of the ball at the hoop. This is a model system to study throwing motions due to its unique position as a pure measure of shooting skill unfettered by defensive pressure and changing shot conditions. Further, there is evidence that even among elite and expert players there is room for marked improvement in free-throw shooting, evidenced by the fact that league-wide average free-throw percentages among professional and collegiate players has remained at a mediocre level of 70-75% for over 50 years. Therefore, free-throw shot training is also a good candidate to specifically study learning within the context of longer delays in feedback (Branch, 2009).

Importantly, we studied the effect of latencies on motor skills in basketball shooting not by increasing feedback latencies, as has been done traditionally, but by eliminating the existing delay with a predictive feedback system to determine if a *reduction* in latency could *improve* free-throw learning. This presents a challenge as it requires using information available early in the shot to accurately predict shot success. As previous work suggests that even small latencies (<100 ms) can impair skill-acquisition, the system would need to produce accurate predictions based on a very short time window of information which places stringent requirements on measurement precision.

We have developed a camera-based shot tracking system that uses early measurements of the ball's motion state after it leaves the shooter's hand to predict the ball's landing position at the rim within 2.5cm and binary shot outcomes with 99% accuracy. Here we use this system to provide shooters with visual and auditory feedback about the future outcome of their shot within just 35 ms of when the ball leaves their hand, essentially eliminating the feedback delay inherent to free throw shooting. This allowed us to study the effect of long feedback delays in motor skill performance and learning.

2.3 Methods

2.3.1 Predictive delay cancellation system

We developed a predictive feedback delay cancellation system for free throw shooting which has the ability to give shooters feedback regarding the accuracy of their shots and its binary success within 35 ms of the ball leaving their hand, reducing the inherent 900-1000ms delay in free-throw shooting by a factor of approximately 25x. An illustration of this system is shown in Fig. 2.1a. We track the ball's motion by capturing images in real-time by a pair of carefully calibrated stereo vision cameras with a combined calibration error of 0.1 pixels running at 80 frames/sec. Synchronously captured images covering a 12.5 ms time window after release of the ball from the shooter's hand transferred to a computer where they are analyzed in MATLAB (Mathworks, Natick, MA) to produce estimates of the ball's initial linear and rotational motion. Using previously fitted coefficients for a polynomial expansion of the ball's initial motion state, a predictive model is used to predict the ball's landing location within 2.5 cm. This predicted landing location is then combined with a classification model which predicts whether the shot will be a direct make based on an estimated success region inside the rim with 99% accuracy. Direct makes are high accuracy shots that quickly and directly pass through the rim with minimal bouncing and rim contact. Indirect makes are low accuracy shots, often less accurate than many misses, which go in by lucky bounce and are thus unrelated to shooting skill. As such, the system reports a refined binary shot outcome of whether the shot was a direct make, which is tightly correlated with shot accuracy. Visual feedback is delivered to the shooter through a 24" 240 Hz refresh rate screen mounted behind the glass

backboard and directly in the shooter's line of sight, and auditory feedback was delivered via a wired speaker placed directly underneath the rim. The total latency of the system, measured as the time between detection of shot release and when the feedback appears on the display screen, was 37 ms.

2.3.2 Participants

14 (2 female) players with organized basketball experience were recruited to participate in the study. None of the shooters reported having any current injuries that would affect their shooting ability. Participants were randomly assigned to either an advanced feedback (AF) group or a non-advanced feedback group (NAF). A total of 7 shooters were assigned to the AF group and 6 to the NAF group. Two participants dropped out of the study due to falling sick with Covid-19. The study protocol was approved by the Harvard University Institutional Review Board, and all participants gave informed consent.

2.3.4 Experimental design

Free-throw shot details

Players were instructed to shoot standard free-throws, which are shots at a rim of height 10 ft from a distance of 15 ft from the backboard (13 ft 9 in from the center of the rim). The rim and backboard used met NCAA regulations for size. An all-black colored regulation 29.5" circumference Men's basketball (Wilson Evolution) was used for shooting, which was modified with painted gray dots to facilitate ball spin tracking. Players were instructed to shoot however they felt comfortable. A rebounder was present throughout the shooting session to collect the ball after every shot so that the shooter did not have to move their feet or chase after the ball.

Training schedule

A total of 3200 shots were record from each player spread across 6 shooting sessions within a two-week period, and participants were only allowed to perform one shooting session in a given day. To mitigate fatigue during shooting sessions, players shot in short 6-8 minute blocks of 50 shots with 2 minutes of rest in between when they were allowed to drink water and sit in a chair. A diagram of the training design is

shown in Fig. 2.1b. On Day 1, participants took 500 normal shots without any additional feedback about shot accuracy from the system as a measure of their baseline. This was followed by 200 shots of training where they received either additional time-advanced feedback (AF) or non-advanced feedback (NAF) on every shot. From Day 2 onwards participants took 500 shots during each session with AF or NAF feedback on every shot. Each session lasted approximately 60-90 minutes depending on the individual players' comfortable speed of shooting. We choose 500 shots per day as a compromise among feasibility of experiment duration, shooter fatigue, and precision in estimating each shooter's true make percentage on a given day (which for an average 70% free throw shooter is expected to give us 2% precision in determining their true shot percentage when calculated as the standard deviation of a binomial distribution with success probability p=0.7).

Visual and auditory feedback

During training blocks players were given additional visual and auditory feedback on every shot. Visual feedback encoding the accuracy of the shot was displayed on a screen mounted behind the glass backboard directly in the shooter's line of sight. Examples of the visual feedback are shown in Fig. 2.1c. The predicted landing location of each shot relative to the center of the direct-make region and normalized by the radius of the region was displayed as a colored dot, overlaid on a white circle representing the direct-make region. The position of the dot encoded the vector error of their shot, while the color indicated the accuracy. Shooters were instructed before training that green dots represent excellent or high accuracy shots which are direct-makes. Yellow dot coloring represents a shot with moderate accuracy that may still go in but mostly by luck. Blue dot coloring represents shots with poor accuracy that are not very unlikely to go in except by luck. Accompanying this visual feedback, an auditory "Super Mario" coin sound was played concurrent with the display of visual feedback for direct makes (green dots), and a "bump" sound similar to the sound of the ball hitting the rim was played for shots of moderate (yellow dots) or poor (blue dots) accuracy.

During shooting, a running count of total makes, total direct makes, and total attempts was displayed on the screen in between shots. For participants in the AF group, predictive visual feedback was delivered for a brief 50 ms "blip" of time at an average latency of 35 ms after releasing the ball, along with simultaneous auditory feedback. The NAF group was also given predictive feedback to maintain consistency in feedback quality across the two groups, though it was delivered concurrent to the ball reaching the rim. None of the participants in either reported having trouble noticing or paying attention to the visual or auditory feedback.

Monetary incentive

Participants were given a base pay of $15/hr. During training, participants were award bonus pay depending on their performance on the current block relative to their performance over the last two blocks.

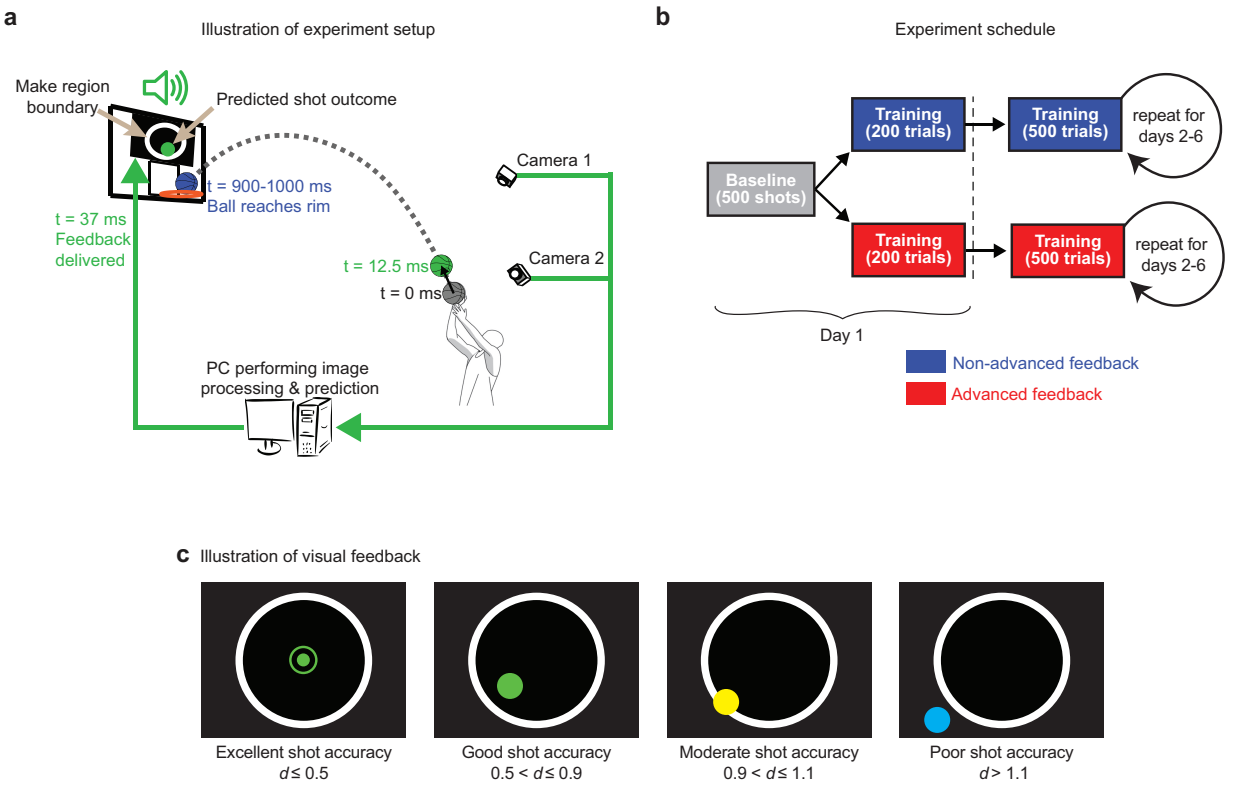

Figure 2.1: Experimental design

(a) Diagram of the predictive delay cancellation system setup. A pair of stereo vision cameras capture 2 image frames of the ball across a 12.5 ms time gap after it is released from the shooter's hand. Images are transferred to a PC which analyzes the images and estimates the initial linear and rotational motion of the ball. This information fed into a model which predicts the landing location of the ball relative to a success region inside the rim to yield both error vector and binary success feedback. This information is encoded visually and auditorily and sent to a display mounted in the shooter's line of sight behind the glass backboard. This feedback is available to the shooter within just 35 ms of the ball's detected release from their hand, circumventing the inherent physical delay in shooting and reducing the feedback latency by a factor of ~25x.

(b) Diagram of experiment design. Players take 500 shots on Day 1 with no additional screen or auditory feedback other than what they would normally see during free throws. Afterwards, they are exposed briefly to the training for 200 shots for their respective group: either the advanced feedback (AF) group where feedback is delivered at t=35 ms, or the non-advanced feedback (NAF) group where feedback is delivered at the time the ball naturally arrives at the rim. In both groups, prediction-based feedback is given to the shooter to keep the quality of feedback consistent across groups. Days 2 through 6 are comprised of 500 shots on each day.

Figure 2.1 (continued)

(c) Example illustrations of screen feedback for different levels of shot accuracy. Participants were shown visual feedback of the vector error of their shot relative to the center of the success region, encoded by the position of a colored dot relative to the center of success region (white). Color of the dot encoded qualitative information about the success of their shot, which was defined by the distance of the shot's predicted landing location relative to the center of the success region, normalized by the success region radius. Shots with excellent accuracy received feedback with a bull's eye target instead of a solid dot.

2.4 Results

Here we studied changes in performance in basketball free-throw showing when the inherent feedback delays in shooting are nearly eliminated in a predictive delay cancellation system, which can provide shooters with time-advanced feedback just 35 ms after they release the ball, reducing the normal 900-1000ms delay by ~25x. We compared the performance for shooters in two groups: an advanced feedback (AF) group which received time advanced feedback 35 ms after releasing the ball, and a non-advanced feedback (NAF) group. We compare performance both in the space of binary success, which we calculate as the percentage of direct-makes, and in a continuous error space, where shot endpoint error is defined as the distance between the shot's landing location at the height of the rim and the center of the success region for direct makes. As previous work has shown that latencies cause decrements to motor learning, we hypothesized that dramatically reducing the delay in the AF group would cause their shooting performance to improve compared to the NAF group.

2.3.1 System feedback quality is consistent across groups

During design of the predictive system, considerable importance was given to improving consistency in predictive performance across shooters. We checked how similar both experimental groups were in terms of the binary and continuous error feedback (Fig. 2.2a). Average classification accuracy for the AF and NAF groups were 1.1 ± 0.1% and 1.3 ± 0.2% [mean ± sem] respectively. For shot endpoint, respective RMS errors were 22.5 ± 1.6 mm and 23.0 ± 2.6 mm [mean ± sem]. Differences between groups for both measures of feedback quality were not significant.

2.3.2 Shooter fatigue is well controlled with experimental design

In our experiment, shooters take 700 shots on Day 1 and 500 shots on Days 2 – 6, where shooters take shots in blocks of 50 shots with 2 minute rest breaks in between. As it was unclear from previous work whether fatigue would play a role in shot performance with this number of shots, we checked whether shooter performance dropped within blocks (Fig. 2.2b, left panel) and within days (Fig. 2.2b, middle panel). We did not see any consistent changes in shooter performance for either within blocks or within days, suggesting fatigue was not a factor in shooter performance. As fatigue could reduce performance towards the end sessions that would have otherwise actually shown improvement, we also estimated fatigue separate from shot performance as a drop in shooting rate. Shooter's average shooting rate did not show any significant changes throughout a shooting session. As we gave shooter's the freedom to shoot at a rate comfortable for themselves, this also suggests that shooter fatigue was well controlled in our experiment design.

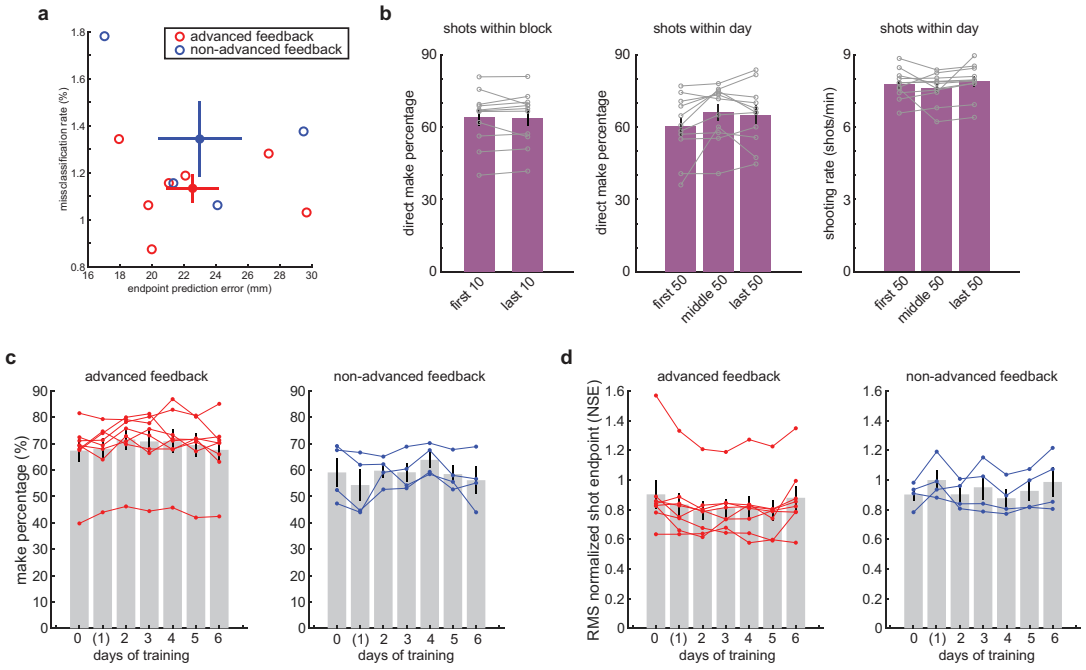

Figure 3.2: Overall performance

(a) Plot comparing prediction quality for AF and NAF participants. Classification accuracy was calculated as the percent of shots whose measured feedback category was "good" or "excellent" but predicted to be "poor", combined with the percent of shots measured to have "poor" accuracy but were predicted to be "good" or "excellent". Shots falling in the "moderate" category accounted for 20% of shooter's shots on average. Classification accuracy is plotted against prediction accuracy in millimeters for individual shooters in the AF (red) and NAF (blue) experimental groups. Overall, system performance was similar across groups.

(b) Fatigue metrics. We limited shot sessions to 500-700 shots in day, where shooters shot in short blocks of 50 shots with 2 min rest breaks in between to limit fatigue. To look for signs of fatigue we compared shot success at the beginning and end of blocks (left panel) and the beginning, middle, and end of sessions (middle panel). We see no significant differences in shooting percentages in these comparisons. As another measure of fatigue we looked at average shooting rate, which also showed no significant changes through a session.

(c-d) Summary of performance metrics throughout the experiment. Binary (**c**) and continuous (**d**) measures of success show similar levels at baseline across both AF and NAF groups. We were not able to detect any significant changes throughout training for either group.

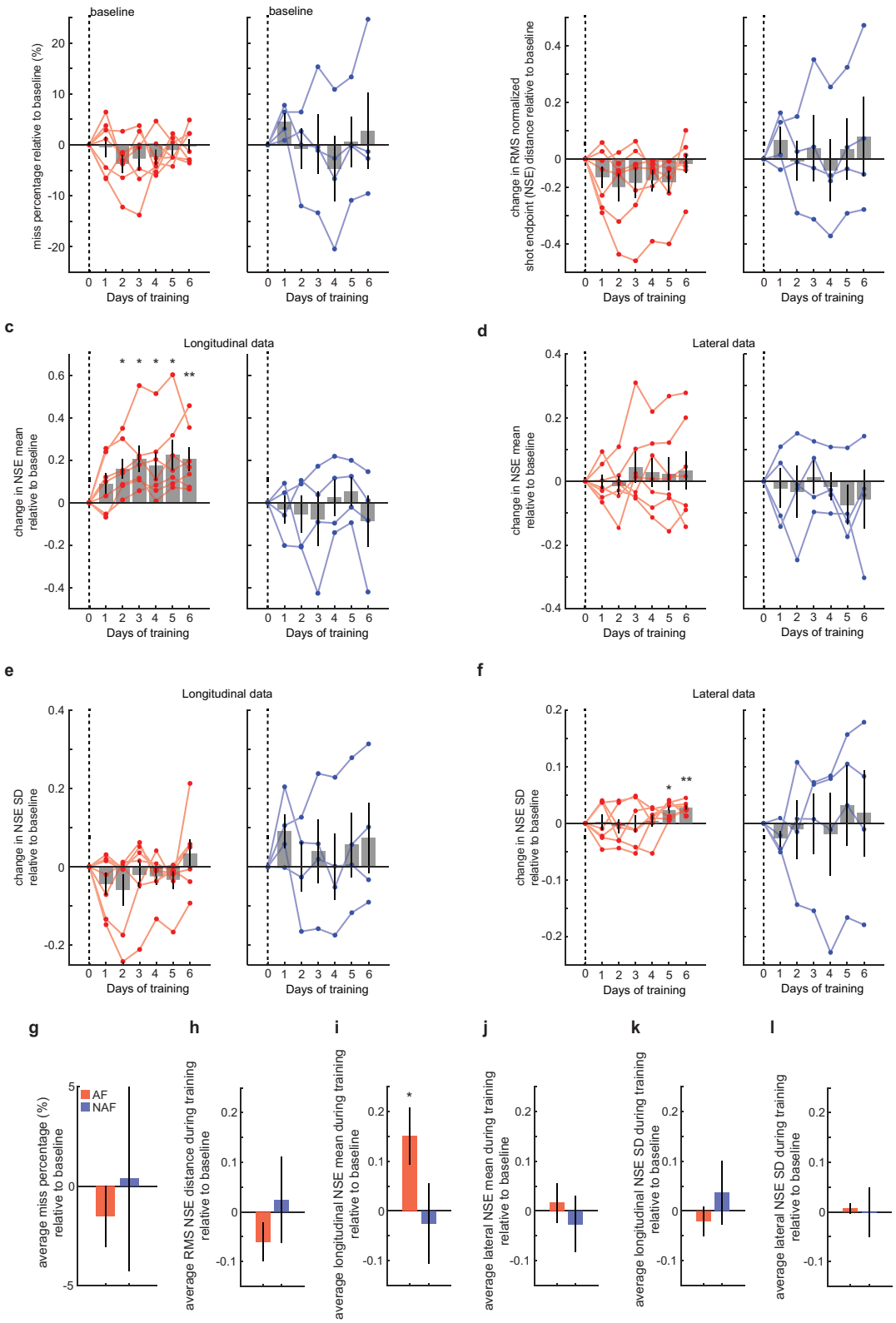

Figure 2.3: Effect of training on shooter performance

Figure 2.3 (continued)

(a-f) Changes relative to baseline for individual shooters are shown in red and blue for shooters in advanced feedback (AF) and non-advanced feedback (NAF) groups respectively, for days 1 through 6 of training. Mean across participants is shown in gray. Note that training on day 1 consists of 200 shots, while training on subsequent days consists of 500 shots. Normalized shot endpoint (NSE) refers to the ball's landing location at the rim relative to the make region center, expressed as a fraction of the make region radius. The lateral direction refers to left-to-right direction across the rim, while longitudinal direction refers to the front-to-back direction across the rim.

(a) Changes in "miss percentage", where we define miss percentage as 100 - direct-make percentage. Plotted are changes in miss percentage relative to baseline, where negative values signify improvement. No significant differences from baseline across subjects were detectable for single days of training.

(b) Changes in RMS normalized shot endpoint (NSE) relative to baseline. NSE is defined as the distance between a shot's landing position relative to the center of the direct-make region, normalized by the diameter of the direct-make region, and so negative values signify improvement. No significant changes consistent across shooters could be detected for single days of training.

(c) Changes in mean longitudinal NSE relative to baseline. AF group shooters showed consistent positive change in longitudinal NSE mean by day 6 of training, while no consistent pattern of change could be seen for NAF shooters. Positive longitudinal shifts are shifts towards the back of the rim. As all shooters had mean longitudinal offsets at baseline that were substantially negative, meaning they consistently shot short of the center of the direct make region, this indicates an improvement.

(d) Changes in mean lateral NSE relative to baseline. As shooters during baseline are on average aligned laterally on with the make region center, any patterned change reflects performance decrement, and no patterned changes could be detected for single day of training.

(e) Changes in longitudinal NSE standard deviation relative to baseline, where negative values indicate improvement. No significant changes could be detected for single days of training.

(f) Changes in lateral NSE standard deviation relative to baseline, where negative values indicate improvement. No significant changes could be detected for single days of training

(g-l) Summary plots of performance metrics relative to baseline, averaged across days of training. Individual shooters performance metrics during training were made relative to their baseline, then data were pooled across shooters and days of training for AF and NAF groups. (**h**) In pooled data, we detect a decrease in RMS NSE for the AF group but not the NAF group, significant at $p<0.001$. Looking at changes in (**i-j**) mean and (**k-l**) standard deviation of NSE reveals that this decrease is mostly driven by improvement in mean longitudinal NSE. Note that we tested (1) whether AF relative to baseline was different from zero, (two tailed ttest) (2) whether NAF relative to baseline was different from zero (two-tailed ttest), and (3) whether AF relative to baseline was different from NAF relative to baseline (unpaired 2-sample ttest). Statistically significant comparison are labelled on the plots with *$p<0.05$, ** $p<0.01$. Non-significant comparisons ($p>0.05$) are not labelled.

2.3.3 Binary and continuous measures of overall shot accuracy do not show significant changes over the course of training

As previous work studying the impact of latencies on motor learning has shown that increased latencies negatively impact performance and the components of motor learning important for complex motor skill acquisition (Brudner et al., 2016; Schween & Hegele, 2017), we hypothesized we would see improvements in overall shooting accuracy in the AF group that were larger than any seen in the NAF group. We calculated changes in binary success rates as the change in percent of shots that were direct makes relative to baseline (Day 1, first 500 shots with no additional screen or auditory feedback) among shooters in both groups (Fig. 2.3a). Neither group showed any significant changes from baseline in their direct make percentage. As shot endpoint error is a more fine-grained measure of shot accuracy, we also calculated changes in RMS make region normalized shot endpoint error across days for individuals in both groups, relative to their baseline (Fig. 2.3b). As the position of the shot endpoint relative to the center of the elliptical make region, normalized shot endpoint can be interpreted as the distance from the make region center expressed as a fraction of the region radius, with shots having a distance less than one being categorized as direct makes. Once again, we did not see any significant changes from baseline shot accuracy in either group. As the RMS shot endpoint error combines information about mean errors with error variance, we thought to see whether either of these components of shot accuracy were changing differentially.

2.3.4 Mean longitudinal distance to the make region center improves only for shooters who receive time-advanced feedback

Interestingly, all shooters displayed a negative mean offset in their longitudinal normalized shot endpoint relative to the center of the make region during baseline shooting, meaning they were consistently shooting short of the make region center. The mean longitudinal offset across shooters was -0.34 ± 0.05 [mean ± sem]. Fig. 2.3c shows changes in longitudinal mean offset for shooters AF and NAF groups (left and right panel, respectively). Across shooters in the NAF group there was no significant difference in

longitudinal mean offset on Day 6 of training relative to baseline (0.03 ± 0.13). However, in the AF feedback group, all shooters showed a positive shift in their mean longitudinal offset as early as Day 2 (0.15 ± 0.04, $p<0.01$) which persisted through Day 6, where the mean offset was 0.19 ± 0.05 with $p<0.01$. As all shooters tended to shoot short of the make region center, this positive shift in mean shot endpoint offset suggests that shooters in the AF group were improving the alignment of their shot endpoint distributions with the make region. As shot endpoint variance tends to be much larger than the mean (5.4 ± 2 [mean ± sem] for ratio of shot endpoint standard deviation vs shot endpoint mean), these shifts would cause small changes in the overall performance metrics that would be difficult to see with only 500 shots in each session.

As shooters in the AF group were shifting their mean longitudinal shot endpoint offsets, we looked to see whether this was accompanied by changes in shot release angle or increases in shot initial speed, both of which could lead to a distribution of shots which land further than baseline on average. While we found no significant change in mean release angle relative to baseline across participants on Day 6 (0.13 ± 0.9 deg, $p=0.9$; -1.4 ±1.3 deg, $p=0.34$ for AF and NAF respectively), we saw a mean increase in shot initial speed for the AF group (6.4 ± 2.2 cm/s, $p<0.05$) that was moderately significant. However, looking at change in initial shot speed within individuals, we found 7 of the 8 participants showed an increase in their release speed that was highly significant ($p<1e-7$ for 7 of 8 AF group individuals; $p<1e-20$ for 5 of 8 AF group individuals). Across shooters in the NAF group we did not see any consistent pattern in changes to their initial shot speed relative to baseline (-2.4 ± 3.7 cm/s, $p=0.53$).

For both AF and NAF groups we saw no consistent changes in mean lateral shot endpoint relative to baseline (Fig. 2.3d). This was unsurprising as shooters were on average laterally aligned with the make region center, with an average lateral endpoint offset of -0.05 ± 0.04, $p=0.23$.

2.3.5 Increased lateral shot endpoint variance accompanies mean longitudinal shifts for AF

We continued to dissect changes in RMS shot endpoint error by examining changes in participant endpoint variance, though given the fact that RMS endpoint errors are dominated by variance compared to offsets we did not expect to see significant differences. Fig 2.2e summarizes the changes in longitudinal shot endpoint standard deviation for days of training relative to baseline for AF and NAF groups (left and right panels, respectively). Neither group showed any significant changes across individuals on any day of training. Fig. 2.3e summarizes changes relative to baseline for lateral shot endpoint standard deviation. While the NAF group showed no significant changes relative to baseline on any day that were consistent across shooters, the AF group showed an average increase in lateral shot endpoint standard deviation that was significant on Day 5 and Day 6 (0.025 ± 0.005, $p<0.01$; 0.031 ± 0.004, $p<0.001$; for Day 5, Day 6 respectively). This was surprising, as an increase in lateral shot endpoint standard deviation would amount to a decrement in performance. It is unsurprising, however, that this decrease in lateral shot endpoint variance did not make significant impact on overall changes in RMS endpoint error shown in Fig. 2.3b, since it amounts to an average increase in lateral shot endpoint variance of 0.4%. This would be expected to only decrease an average shooter's direct make percentage by 0.25%.

2.5 Discussion

2.5.1 Summary

Here we examined the effect of eliminating feedback delays on performance in basketball free-throw shooting. Previous work has indicated that latencies between motion and feedback attenuate motor learning, albeit while studying tasks of lower complexity and much smaller feedback delay latencies than those seen in basketball shooting and other throwing tasks (Brudner et al., 2016; Schween & Hegele, 2017). We implemented a predictive delay cancellation system that can reduce the normally 900-1000 ms feedback delay in basketball by a factor of ~25x, reducing it to 35 ms. To examine the effect of removing the feedback latency, collected 3,200 shots from 14 experienced basketball players, where 8 participants were trained on shots with time-advanced predictive feedback just 35 ms after releasing the ball, and 6 participants were trained on shots with similar predictive feedback that was instead shown concurrent with the ball reaching the hoop.

In summary, though we did not see improvements in overall shooting percentage or overall RMS shot endpoint errors for the AF group compared to the NAF group, we did see significant changes specific to the AF group when dissecting shot endpoint errors into changes in mean offsets and changes in variability. While these changes do not amount to large changes in overall performance, they indicate promise for future work with improved experimental design.

2.5.2 Lack of overall improvement from predictive feedback in AF group

Unfortunately, we did not see any significant changes in shooter's overall direct make percentages or RMS shot endpoint errors. We were expected that with such dramatic reduction in feedback latency for the AF group we would see an increase in direct make percentage or a decrease in RMS shot endpoint errors. This could be due to multiple reasons unrelated to whether time-advanced feedback is truly helpful for improving performance. First, it is possible shooters are very sensitive to misclassification and prediction errors and quickly lose faith in the feedback shown on the display any time a misclassification

occurs, despite the fact that most shooters experienced a misclassification on only 1 shot in 100 on average. It is also possible that shooters have strong inertia towards changing their shooting style, even when presented with useful information that nudges them to do otherwise. Most of the shooters in our experiment have been playing basketball regularly for several years, and so it would be reasonable to think that changing their shooting style within five days of training would be difficult. In future work, it will be interesting to train naïve participants with no prior experience playing basketball. This would likely increase the magnitude of the learning signal for two reasons. First, these participants would have much room for improvement. Second, these naïve participants are less likely to have inertia in their initial shooting style, and so that may also increase their rate of learning. Given feasible limits on the length of shooting sessions, this has the potential to allow easier detection of the effects of delay cancellation

In our experiment, shooters in the AF group received feedback 35 ms after releasing the ball. However, the flight of the ball and its behavior at the rim, which is the type of feedback they have been trained on for many years, is still readily available to them on every shot. The physical outcome of the shot, and their understanding of those outcomes from years of shooting, may be another interference with their use of the available early feedback. It is possible that a paradigm which prevents the physical outcomes of shots from being available to shooters would encourage more potent use of the time-advanced feedback.

2.5.3 Existing mean longitudinal endpoint offsets in baseline towards the front of the rim likely reflect poor estimation of true longitudinal shot accuracy by shooters

The classification model implemented in our system calculates a direct make region specific for each shot depending on the shot's approach angle at the rim. This make region critically separates direct makes, which are a clarified subset of made shots which are strongly correlated with actual shot accuracy, from other shots. Indirect makes, however, are low accuracy shots that occur in a mostly isolated region towards the front of the rim, and end up going through the rim by lucky bounce. Often, a shot similar to an indirect make with slightly improved accuracy would actually generate a complete miss, and so the region of shot endpoints relative to the rim position containing most indirect makes creates a somewhat

discontinuous error surface for players to follow while learning to shoot. Analysis of shots detected as direct makes and indirect makes shows that the percent of indirect makes is largely unrelated to shooting skill, and so following this reward signal would be unhelpful.

However, shooters are not aware of this true mapping of direct and indirect makes. Additionally, even if they were, shooters must estimate longitudinal errors in the shots landing point from a vantage point that makes it difficult to see small changes in the ball's landing point. In contrast, shooters have a good vantage point of lateral errors, as they can easily see if the ball is landing on the left or right side of the rim. The behavior of the ball as it lands on the rim further complicates estimation of precise shot accuracy, as shots hitting the front of the rim miss less dramatically than shots with similar accuracy that hit the back of the rim. It is likely that shooters would see indirect makes as being shots of accuracy just shy of direct makes, which we know is not true. Further, it is possible that shooters' estimated accuracy of misses off the back rim, which bounce much farther and more dramatically than a shot of similar accuracy off the front rim, would be unfairly colored by the resulting bounce. Together, it is likely that shooter's initial mean longitudinal offsets, which are short of the direct make region center, reflect an inclusion of the region of indirect makes.

2.5.4 Improvement in mean longitudinal endpoint offset for AF group shooters suggests possible improvement from elimination of feedback delays

The changes mean offset for the AF group towards better alignment with the direct make region is a minor improvement, as it does not effectively changes shooters' direct make percentages or RMS endpoint errors. However, it does represent an improvement that is specific to shooters receiving time-advanced feedback. Though it is not strong evidence for improvement from predictive delay cancellation, it suggests that improvements to the current experimental design could show much stronger effects.

2.5.5 Increase in lateral endpoint variability in the AF group

Changes in mean longitudinal offset in the AF feedback by Day 6 were accompanied by a small but statistically significant increase in lateral shot endpoint variability. This is surprising, given that increased variability represents a decrease in overall performance. However, this decrement in performance is likely mostly undetectable to shooters, as the average increase in lateral endpoint variability would only result in a 0.25% decrease in direct make percentage for an average 70% free throw shooter. One possible explanation is that the increase in initial ball speed displayed by shooters in the AF group, in service of the changes in mean longitudinal offset that we see for the same group, increase some small amount of lateral endpoint noise as signal-dependent noise.

2.6 References

Branch, J. (2009, March 3). For Free Throws, 50 Years of Practice Is No Help. *New York Times*.

Brudner, S. N., Kethidi, N., Graeupner, D., Ivry, R. B., Taylor, J. A., & Taylor, J. A. (2016). Delayed feedback during sensorimotor learning selectively disrupts adaptation but not strategy use. *J Neurophysiol*, *115*, 1499–1511. https://doi.org/10.1152/jn.00066.2015.-In

Kehoe, E. J., Ludvig, E. A., & Sutton, R. S. (2009). Magnitude and Timing of Conditioned Responses in Delay and Trace Classical Conditioning of the Nictitating Membrane Response of the Rabbit (Oryctolagus cuniculus). *Behavioral Neuroscience*, *123*(5), 1095–1101. https://doi.org/10.1037/a0017112

Kumcu, A., Vermeulen, L., Elprama, S. A., Duysburgh, P., Platiša, L., van Nieuwenhove, Y., van de Winkel, N., Jacobs, A., van Looy, J., & Philips, W. (2017). Effect of video lag on laparoscopic surgery: correlation between performance and usability at low latencies. *International Journal of Medical Robotics and Computer Assisted Surgery*, *13*(2). https://doi.org/10.1002/rcs.1758

Najafi, F., & Medina, J. F. (2013). Beyond "all-or-nothing" climbing fibers: Graded representation of teaching signals in Purkinje cells. *Frontiers in Neural Circuits*, *JUL*. https://doi.org/10.3389/fncir.2013.00115

Ohmae, S., & Medina, J. F. (2015). Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. *Nature Neuroscience*, *18*(12), 1798–1803. https://doi.org/10.1038/nn.4167

Roach, N. T., Venkadesan, M., Rainbow, M. J., & Lieberman, D. E. (2013). Elastic energy storage in the shoulder and the evolution of high-speed throwing in Homo. *Nature*, *498*(7455), 483–486. https://doi.org/10.1038/nature12267

Schween, R., & Hegele, M. (2017). Feedback delay attenuates implicit but facilitates explicit adjustments to a visuomotor rotation. *Neurobiology of Learning and Memory*, *140*, 124–133. https://doi.org/10.1016/j.nlm.2017.02.015

Smith, M. C., Coleman, S. R., & Gormezano, I. (1969). Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous, and forward CS-US intervals. *Journal of Comparatvie and Physiological Psychology*, *69*(2), 226–231.

Suvrathan, A., Payne, H. L., & Raymond, J. L. (2016). Timing Rules for Synaptic Plasticity Matched to Behavioral Function. *Neuron*, *92*(5), 959–967. https://doi.org/10.1016/j.neuron.2016.10.022

Young, R. W. (2003). Evolution of the human hand: the role of throwing and clubbing. In *J. Anat* (Vol. 202).

Chapter 3: Detecting streaky performance in free-throw shooting in individual players

3.1 Abstract

On one hand, it is well known that there are continuous drifting fluctuations in biases that effect motor performance, and that it has even been demonstrated that these fluctuations can arise from slow drifts in neural activity in motor cortical areas (Chaisanguanthum et al., 2014). On the other hand, motor variability has been shown to change with a gradual improvement during long-term learning (Cohen & Sternad, 2009; Sternad, 2018), and to change rather abruptly when going from practice to performance states (Kojima & Doupe, 2011). However, it is not known whether continuous drifting fluctuations in motor variability, like those readily observed for biases in motor performance, occur as well. In fact, one instantiation of this question, whether the "hot hand" exists, has been vigorously debated for nearly 4 decades.

The “hot hand” phenomenon refers to a transient increase in ability levels in skilled movements, a term most often used in basketball in referring to a player whose current shooting ability is above her normal level. Since increases in shooting ability require improved precision and thus reduced motor variability, a hot hand would rely on a transient change in motor variability. However, whether the hot hand truly exists, i.e. whether observed increases in shooting accuracy reflect temporary improvements in ability rather than pure luck (statistical fluctuations in shot outcome without an underlying change in ability), is not clear. Its existence has been hotly contested since a controversial study staunchly claimed the hot hand to be a cognitive illusion, a widespread belief among fans, players, and coaches arising only from the human tendency to overestimate the strength of patterns that we observe. Since then, studies attempting to detect whether streaky performances arise from trivial statistical fluctuations versus from true excursions in the underlying ability level have struggled to provide definitive answers. In particular, these studies have struggled to find datasets and methods which provide statistical power requisite to test a key element of the hot hand theory, which is that streaky performance, beyond that expected from random chance, occurs in individual players. Here we bring both more powerful analyses and more powerful data to bear on this problem and provide a clear answer. We introduce a novel method for

detecting streaky performance beyond random chance based on gauging the ability to predict future performance in shot sequences. We combine this with a continuous measurement of shot accuracy rather than the discrete (make/miss) measurement that has traditionally been used, which facilitates detection of streakiness in individuals with an overall 10 to 14 fold reduction in the sample size required to achieve good statistical power compared to existing methods. We analyze a free throw shooting dataset from 16 individuals with 3,200 shots each and find that 11 of 16 individuals display clear evidence of underlying performance fluctuations, with an estimate for the standard deviation of excursions of +/- 8.6% on average across individuals, decisively showing that performance fluctuations exists and is a widespread phenomenon among shooters.

3.2 Introduction

On a Friday evening in 2015, playing against his team's local rivals the Sacramento Kings, Klay Thompson of the NBA-champion Golden State Warriors pulled off one of the greatest offensive feats ever seen in basketball. Thompson scored 37 points in just 10 minutes of play while making all 13 of his attempted shots, setting a National Basketball Association record for most points in a quarter (Haberstroh, 2017; "Klay Thompson Sets NBA Record with 37 Points in a Quarter," 2015) . For perspective, that was more than a third of the average points scored by an entire NBA team per game in the 2014-15 season (https://www.basketball-reference.come) .

Surveys of players and fans alike (Avugos et al., 2013; Gilovich et al., 1985) suggest that many people would claim Thompson's performance is an example of a "hot hand", referring to an individual player's temporarily elevated ability to make shots beyond what is expected from chance given their normal ability level. However, a simple analysis of this 13 for 13 streak reveals that it would have an unremarkable 8% chance of occurring simply by chance at some point during Thompson's 2015 season given the 1386 shots he attempted and the 47% shooting accuracy he displayed over the course of the season. In line with this observation, in a famous study published in 1985, Gilovich et al reported that statistical evidence for the hot hand could not be found, and suggested that the widespread in belief in its existence was the result of a cognitive fallacy born of the human tendency to see non-random patterns in small samples of random sequences (Gilovich et al., 1985). The results of Gilovich et al were highly controversial, however, and prompted a number of subsequent investigations into whether the hot hand exists, with studies arguing both for and against its existence (Arkes, 2010, 2013; Avugos et al., 2013; Bar-Eli et al., 2006; Lantis & Nesson, 2021; Miller & Sanjurjo, 2018, 2019, 2021; Morgulev et al., 2020; Pelechrinis & Winston, 2022).

Close examination of the studies examining the hot hand hypothesis, (henceforth referred to as performance fluctuations) reveals methodological issues that hinder the interpretability of their respective claims, however. These issues can be appreciated by separating previous work into two primary approaches to obtaining datasets of shot sequences. In one approach, shooters take several shots in a controlled setting, outside of games, while researchers record shot outcomes (Avugos et al., 2013; Gilovich et al., 1985; Jagacinski et al., 1979; Miller & Sanjurjo, 2014; Wardrop, 1999), and has led to datasets too small to reliably detect fluctuations in performance. And while more recent work still using this approach has attempted to circumvent this issue by computing these fluctuations as an increase in make percentage for shots following 3 makes compared to the sequence average (Miller & Sanjurjo, 2014), statistical confidence in detected streakiness was moderate, and only 2/14 participants were found to display performance fluctuations using this metric.

The second approach is to obtain in-game shooting datasets containing sequences of shots from individual players. This approach benefits from getting data across many more individuals, however, use of in-game data requires methods to control for varying game situations. Accordingly, some studies have limited analyses only to free-throws, which occur during pauses, are free of defenders, and are consistently taken from the same distance to the hoop (Arkes, 2010, 2013; Lantis & Nesson, 2021; Morgulev et al., 2020; Yaari & Eisenmann, 2011). Unfortunately, such a limitation results in a significant reduction of the data available within individuals, prompting previous studies to *combine* data across individuals. However, doing so assumes that the effect size is consistent across all individuals, which is highly unlikely, and thus underestimates the true effect size. Given this issue, it is unsurprising that one study found no evidence for performance fluctuations using this approach (Lantis & Nesson, 2021; Morgulev et al., 2020), while others that report evidence for such fluctuations do so with moderate to weak statistical confidence and small estimated effect sizes (<5%) (Arkes, 2010, 2013; Yaari & Eisenmann, 2011). The recent availability of more detailed in-game statistics such as distance of each shot, distance of the closest

defender, identity of the defender, or number of dribbles before each shot, has facilitated more detailed models for controlling for shot difficulty (Pelechrinis & Winston, 2022). However detection of performance fluctuation in individuals was still weak.

The inconsistent reports of the studies examining performance fluctuations, as outlined above, stem from weak statistical power, especially within individual shooters, which can be attributed to both small datasets and low-powered metrics. To resolve this, here we introduce a novel and robust metric for detecting streaky performance among shot sequences from individual shooters, termed the performance prediction coefficient, which we combine with continuous-valued shot outcomes measured using a stereo vision ball-tracking system at the rim. This affords an estimated 10-14x increase in statistical power over current metrics. In shot sequences containing 3,200 tracked shots from 16 shooters, which offer a dataset far larger than previous work, we then use this metric on sequences of continuous-shot outcomes to provide incredibly strong evidence of performance fluctuations among individual shooters.

3.3 Results

An important feature of the hot hand theory is that individual shooters are expected to go through periods of better and worse performance. As previous work has struggled to detect performance fluctuations individual shooters due to a lack of statistical power, we investigated a new method for detection of streaky performance which measures the relationship between performance on a given shot and the average performance in a window of immediately previous shots. In particular, we were interested in whether the statistical power afforded by this method was sufficient to detect performance fluctuations in individual shooters. To test our method, we collected 3,200 shots over six shooting sessions from 16 individuals while recording their shot outcomes using a stereo-camera ball tracking system, and used this data to develop more precise binary and continuous definitions of shot outcome.

3.3.1 Measuring shot outcomes

Figure 3.1a shows a diagram of the stereo-vision setup used to measure the ball's approaching trajectory and whether the was a make or miss. Shot endpoint, a continuous-valued measure of shot accuracy, was calculated as the projected intersection point of the ball's trajectory and a horizontal plane at the rim height (Fig. 3.1a). We subcategorized made shots into indirect makes, which are lower accuracy shots that go in after multiple bounces on the rim, and direct makes, which are high accuracy shots that pass through the rim with minimal to no rim contact. Across participants, indirect makes were found to account for 8.2%±0.4% of all shots and 12.2%±0.8% of made shots, and the average direct-make percentage was 61.2%±3.3% [mean±sem]. We designed a classification model which uses an elliptical success region boundary in the space of shot endpoints to separate direct makes from other shots. Additionally, we fit parameters which modulate the geometry of this "direct-make region" based on the approaching angle of the ball, since steeper approach angles increase the effective rim size for the incoming ball (Fig. 3.1b). This model appropriately enlarges the direct-make region with higher approach angles, with an increase in make region area of ~40% across the 5^{th} and 95^{th} percentiles of approach angle. Remapping the shot endpoints for direct makes, indirect makes, and misses in coordinates normalized by the make region for individual shots (Fig. 3.1c) reveals that indirect makes occur in an isolated region of low accuracy shots towards the front of the rim, with accuracies lower than many pure misses. Considering these shots to be an impurity existing in the traditional discrete make/miss definition of success whose prevalence is poorly related to shot accuracy, we use whether a shot was a direct make as the discrete indicator of shot success for subsequent analyses.

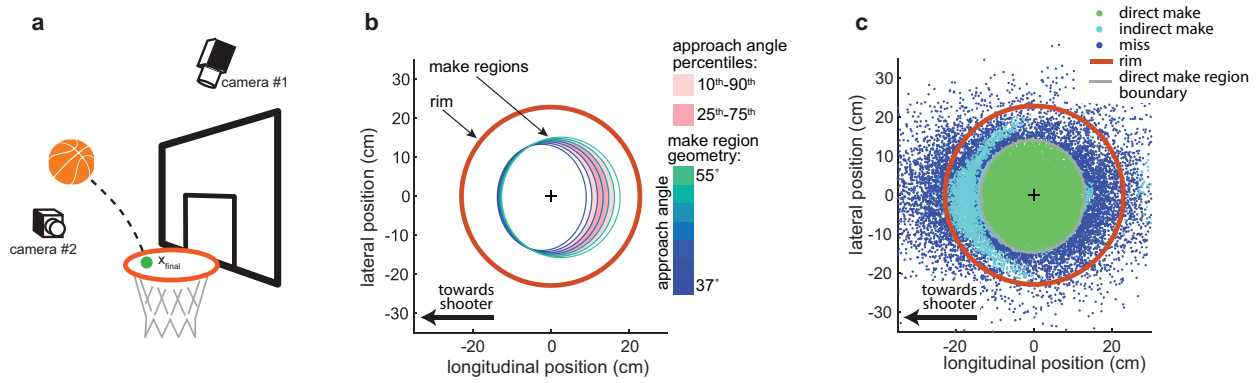

Figure 3.1: A system for precise measurement and classification of shot outcome

(a) Diagram of 2-camera stereo vision system for tracking approaching ball trajectory and precise shot endpoint relative to the rim position. Ball position was tracked with an estimated precision of ~80 µm. We separated made shots tracked by the system based on whether they quickly fall through the rim with minimal contact (direct makes) or if they fall through after multiple bounces (indirect makes).

(b) Diagram demonstrating the relationship between an approaching ball's trajectory and geometry of the direct make region. Increasing approach angles, defined as the angle between the ball's trajectory and the horizontal plane, results in wider direct make regions. Shaded pink regions indicate the general spread of make region geometry experienced by shooters in our experiment, with the middle 50% and middle 80% of make regions shown in dark and light pink respectively.

(c) Map of direct makes (green dots), indirect makes (cyan dots), and misses (blue dots) as a function of longitudinal and lateral positions of the shot endpoint. Here shot endpoint is defined as the location in the horizontal plane at the rim height (10 ft, 3.05 m) that the ball's trajectory would intersect if not obstructed by rim or backboard contact. We used this map to estimate an elliptical "direct make region" (gray) which can separate direct makes from misses and indirect makes with 99% accuracy.

3.3.2 Streak detection in simulated sequences of shots

Detecting performance fluctuations in sequences of shots requires a metric that can estimate a players current performance state, or ability level, in a local portion of the sequence and then seeing to what degree this state can explain performance on subsequent shots. Previous shooting studies have used streaks of 1 to 4 makes as binary estimators of the performance state being "hot" or "cold" (Arkes, 2010, 2013; Avugos et al., 2013; Gilovich et al., 1985; Lantis & Nesson, 2021; Miller & Sanjurjo, 2014, 2018; Morgulev et al., 2020; Pelechrinis & Winston, 2022; Yaari & Eisenmann, 2011). However, for a 70% shooter this is expected to render approximately 75% of the sequence unusable for estimating the make probability of shots during the hot state. We instead choose to estimate the shooter's local performance state as their average performance over a w length window. We then detected performance fluctuation as the slope of a regression, where the outcomes of shots are regressed onto average performance over the immediately previous w shots. We term the slope of this regression the performance prediction coefficient.

We first applied this method to simulated sequences of discrete and continuous shot outcomes with varying degrees of performance fluctuation to evaluate its efficacy in detection. A visualization of the performance prediction coefficient calculated with different window lengths w is shown in Fig. 3.2 a & b for simulated streaky sequences of discrete shot outcomes (Fig. 3.2a) and [matched] continuous shot outcomes based on shot endpoint error (Fig. 3.2b) (simulated sequences had moderately sized excursions in performance state equaling ±10% when expressed changes in make percentage). In discrete shot outcomes, the average performance of w previous shots is naturally binned, as the average over w can only take on fractional values. For discrete shot outcomes the w=1 performance prediction coefficient is equivalent to a metric calculating the difference in make percentage between shots following makes and shots following misses. In both continuous and discrete shot outcomes, the relationship between current and previous average performance becomes steeper as w is increased. Visual comparison of the fitted

slopes (black lines) on discrete and continuous shot outcome sequences shows that these slopes are larger in continuous shot outcomes, indicating better signal to noise ratio.

For discrete and continuous shot outcomes we compared the relationship between performance prediction coefficient and excursion size, which is the estimated amplitude of changes in underlying ability level expressed in the units of make percentage. We used two types of processes to model the underlying temporal changes in ability level. The first is a random jump process in which a shooter's underlying ability randomly switches between hot and cold states with some fixed transition probability (Fig. 3.2c, left panel). The second is a tethered random walk process where a shooter's ability level is allowed to randomly drift from one trial to the next (Fig. 3.2c, right panel). We compared estimates of performance prediction coefficient and excursion size for a 70% shooter using these two underlying processes for a range of window sizes and excursion sizes (Fig. 3.2d). For both discrete (left panel) and continuous (right panel) shot outcomes, performance prediction coefficient increased monotonically with excursion size and window size, with tightly similar results for random jump and random walk process based simulations. The effect of increasing window size on the size of the performance prediction coefficient was dramatic: for discrete shot outcomes, performance prediction coefficients for w=10 and w=20 were a factor 8x and 13x higher, respectively, compared to w=1, and a factor 5x and 3x higher compared to w=3. For continuous shot outcomes the improvement in performance prediction coefficient from using larger window sizes was also large, with a factor 6x and 9x improvement for w=10 and w=20 compared to w=1, respectively; compared to w=3, the respective improvements for w=10 and w=20 were factors of 2.5x and 3.5x. However, comparing continuous shot outcome against discrete shot outcome, performance prediction coefficients were all approximately 2x larger in continuous shot outcomes. Together, this indicates that estimating performance prediction coefficient on sequences of continuous shot outcome using larger window sizes dramatically increases the statistical power of the method.

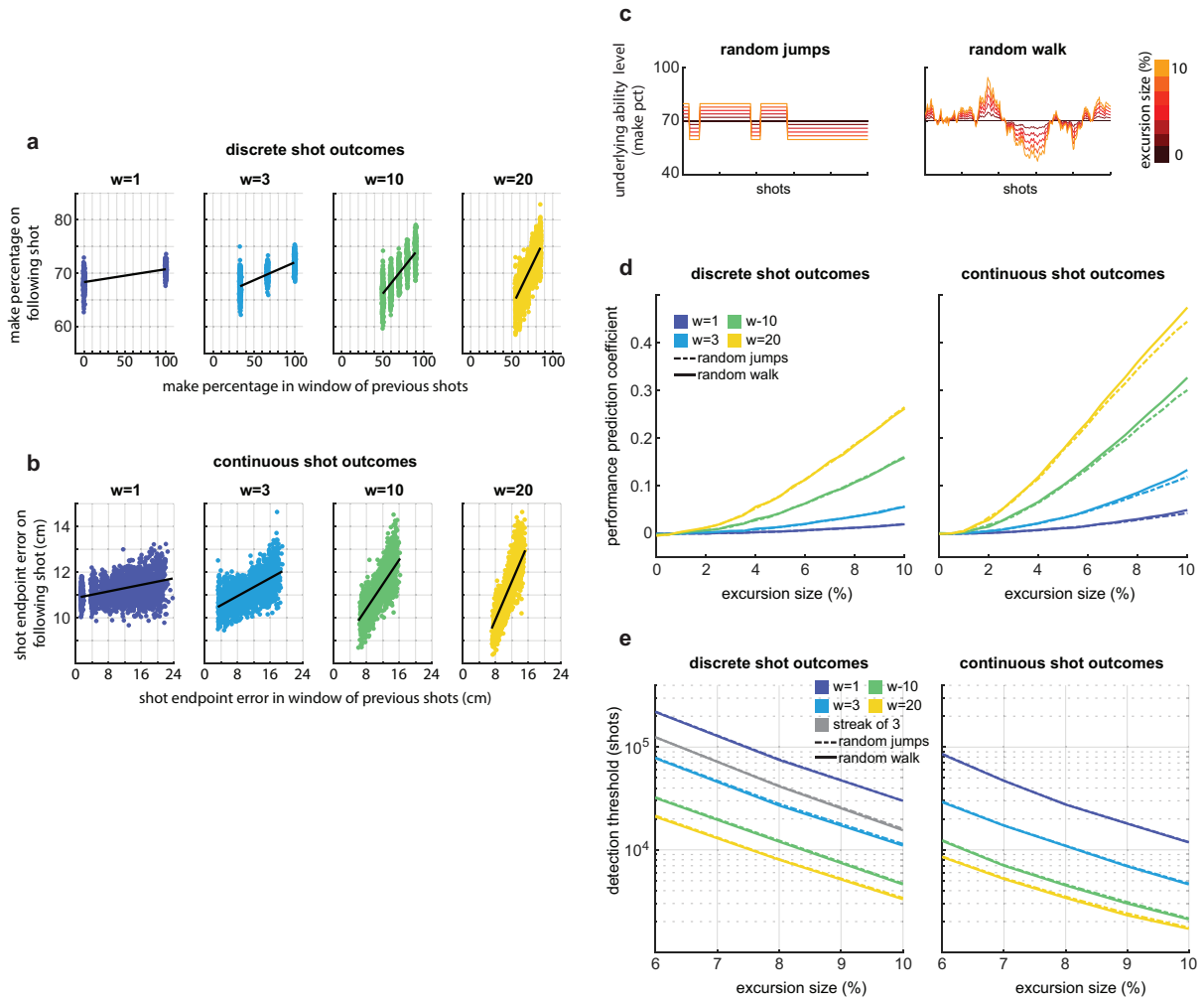

Figure 3.2: Detecting and quantifying fluctuating performance in basketball shooting

We detect performance fluctuations by measuring a performance prediction coefficient, which we define as the strength of the relationship between predicted performance, based on a window of the *w* most recent previous shots, and performance on the subsequent shot.

(a-b) Straight line fits (black) to the relationship between windowed recent performance vs subsequent shot performance (colored dots) for 300 simulated sequences of moderately streaky shots (excursion size of 10deg) when performance is measured based on (**a**) discrete make/miss information, or (**b**) continuous shot endpoint error. Shot endpoint error was quantified as the distance between each shot's endpoint location and the center of direct make region. (For visualization purposes only, data in **b** were binned, with bins containing less than 5% of the data not shown in **a** and **b**, and horizontal visual jitter was added to data in **a**).

Figure 3.2 (continued)

(c) During simulation we considered two types of time-varying patterns for how shot accuracy might evolve: a random jump model, where a shooter's expected make percentage randomly switches between hot and cold states, and a tethered random walk process, where a shooter's expected make percentage is allowed to drift randomly from one trial to the next, both for a 70% shooter.

(d) Simulated relationships between excursion size and performance prediction coefficient for different window sizes show that prediction coefficients increase with excursion size and with window size for a given excursion size, with similar results for the random jump and random walk patterns. A comparison between the left and right panels shows markedly increased performance prediction coefficients for continuous shot outcomes compared to discrete shot outcomes for the same excursion size and window size.

(e) To quantify statistical power, we determined the number of shots required to detect fluctuation at $p<0.01$ significance in 90% of simulations as a function of excursion size for different analysis methods in binary (left) and continuous (right) shot outcomes. Colored lines correspond to different window sizes used in estimation of the performance prediction coefficient. For binary shot outcomes, dark blue lines corresponding to window size $w=1$ are equivalent to a commonly used metric calculating the increase in make probability following a streak of 1 (Gilovich et al., 1985). Gray lines correspond to a streak metric calculated as the difference between make percentage following streaks of 3 makes and overall sequence average (Miller & Sanjurjo, 2014). Estimates of statistical power were similar across simulations based on random jumps (dashed) or random walk (solid).

To quantify the statistical power of our method we calculated in simulation a detection threshold for different excursion sizes, which we defined as the shot sequence length required to detect performance fluctuation in an individual shooter with $p<0.01$ significance in 90% of simulations (Fig. 3.2e). Strikingly, the number of shots required in discrete shot outcomes for window sizes of 1 and 3 shots in the range of excursions tested (up to 10%) is more than the number of shots available for individuals in any dataset to date. For comparison we have included the commonly used metric comparing the make probability following 3 makes to the sequence average (gray) (Miller & Sanjurjo, 2014). Use of this metric is an improvement over w=1, though it is weaker than w=3, requiring ~50% more shots to reach the detection limit. This indicates that previous studies are severely underpowered and have greatly underestimated the effect sizes of performance fluctuations.

Given that the estimated performance prediction coefficient was seen to be larger for larger window sizes, we expected that the number of shots required to reach the detection threshold would decrease with increasing window size. This is indeed the case, as the number of shots required in discrete shot outcomes for a 10 deg excursion size dramatically decreases from 30,000 shots for w=1 to 11,000 for w=3, 11,000 for w=10, and 3,400 for w=20, indicating almost an order of magnitude decrease in the number of shots required which is consistent with the increase in performance prediction coefficient seen in Fig. 3.2d. Similarly to the 2x improvement in measured performance prediction coefficient between discrete and continuous shot outcomes, we also see a 2x improvement in detection thresholds when comparing across discrete and continuous shot outcomes. For of w=10 and w=20, this analysis indicates that performance fluctuation can reliably be detected in individuals with true excursions of 10% with datasets containing just 2,100 and 1,700 shots, respectively, which is a much more tractable number of shots to collect from individual shooters and is of similar order compared to previously collected datasets (Miller & Sanjurjo, 2014).

3.3.3 Detection of performance fluctuation in individual shooters

We collected 3,200 shots over 6 shooting sessions from 16 shooters and analyzed detectable performance fluctuations in their sequences of shot outcomes. Fig 3.3a shows moving average filtered traces of shot endpoint error for 4 example subjects. We measured their performance prediction coefficients using window sizes of 1, 3, 10, and 20 shots while respecting boundaries between sessions. Visual demonstration of performance prediction coefficient estimation corresponding to Fig. 3.3a are shows in Fig. 3.3b. Across individuals, we saw strong evidence for performance fluctuation across a range of window sizes for both discrete (0.019±0.006, 0.04±0.01, 0.10±0.02, 0.12±0.04; p=0.004, p=0.0005, p=0.0004, p=0.003; for w=(1, 3, 10, 20); t-test) and continuous shot outcomes (0.041±0.009, 0.12±0.02, 0.24±0.03, 0.31±0.05; p=0.0004, p=1.3x10^-5, p=6.5x10^-6, p=1.1x10^-5; for w=(1, 3, 10, 20); t-test). These results corroborate our simulation findings well, as average estimates of performance prediction coefficient increased with increasing window size in both discrete and continuous shot outcomes and continuous shot outcomes yield performance prediction coefficients that are ~2x larger than those in discrete shot outcomes (see Fig. 3.2 d & e).

As average estimates of performance prediction coefficient across individuals demonstrated that performance is strongly present at the group level, we next looked at whether we could detect fluctuating performance within individuals. Strikingly, we found incredibly strong evidence that the majority of individual shooters displayed fluctuations, as 11 of 16 shooters had measured performance prediction coefficients significant at p<0.01 level when calculated using w=20 on continuous shot outcomes. Of these 11 shooters, 8 displayed fluctuations in performance at significance level p<0.001 and 5 displayed fluctuation with p<1x10^-9. Additionally, we also found incredibly strong evidence of fluctuating performance in aggregate data when controlling for individual shooter ability level, where the performance prediction coefficient was estimated to be 18.3±0.03 with p=7x10^-159 [mean±95% CI]. A plot of the statistical strength of individual estimates of performance prediction coefficients calculating with different window sizes is shown in Fig. 3.3e, which summarizes the striking confidence in detecting streaky performance within individual shooters when measuring performance prediction coefficients on

continuous shot outcomes with window sizes of 10 or 20 shots. Note that, to visualize these data, we have resorted to plotting $-\log_{10}(p)$ *on a log scale*, which is indicative of the statistical power of the method.

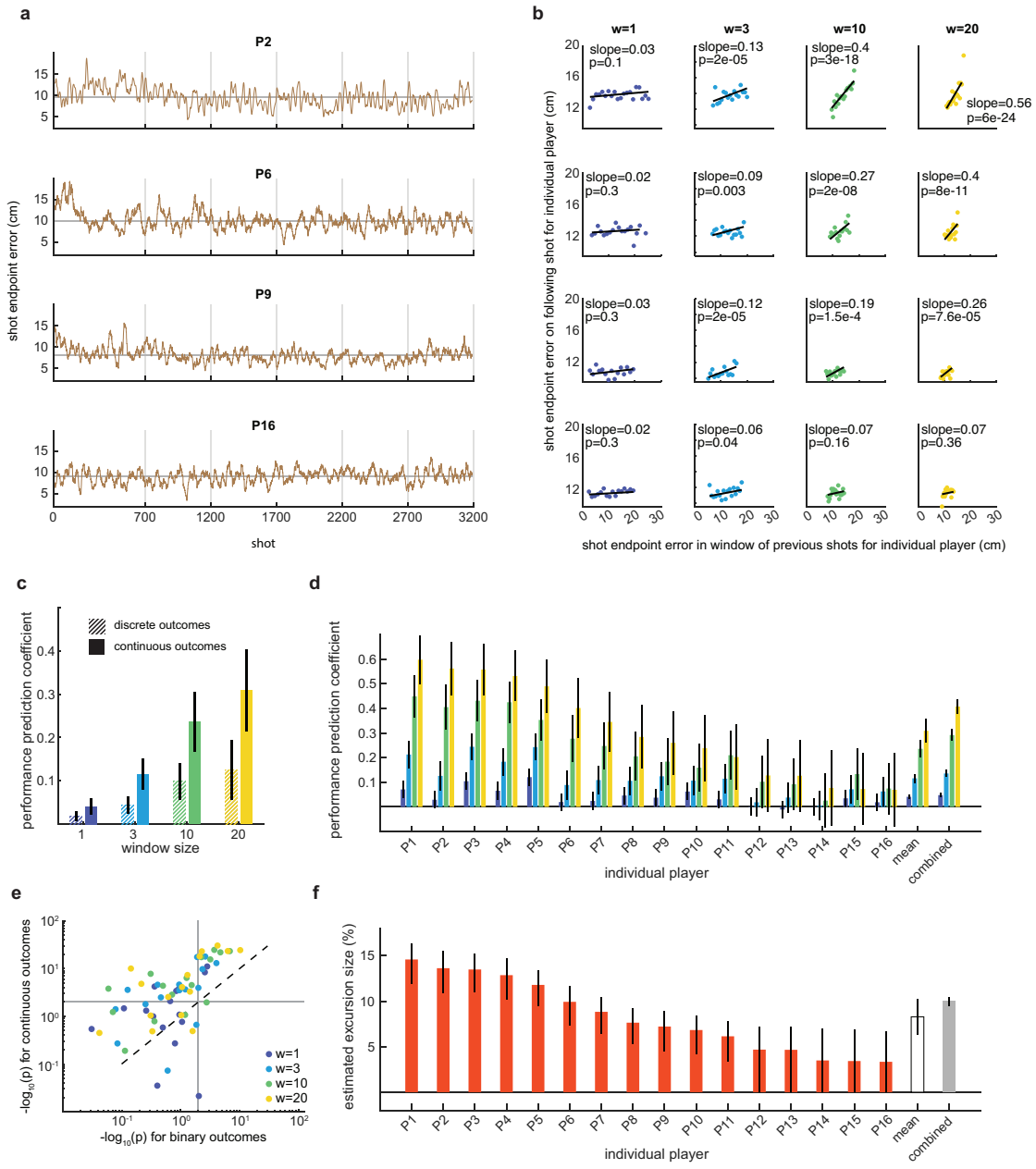

Figure 3.3: Measuring performance prediction coefficients on shooting data decisively reveals fluctuating performance among several individuals and at the group level.

3,200 shots across six days of shooting were recorded from 16 active club team or recreational level basketball players. Performance prediction coefficients were measured for each player's sequence of shot outcomes assuming streaks cannot carry over across days

(a-b) Data for four example shooters shown in descending order of performance prediction coefficient.

Figure 3.3 (continued)

(a) Sequences of measured shot endpoint error (brown) for four example shooters show temporal changes in shot accuracy relative to their overall mean (black). Shot endpoint error was filtered with a moving average of window size 50 to aid visualization. Gray lines indicate boundaries between shooting sessions.

(b) Plots visualizing estimation of performance prediction coefficient with different window sizes in sequences of continuous shot outcome from example shooters. Rows of panels correspond to individual shot endpoint error sequences shown in **b**, and data was binned for visualization purposes. Performance prediction coefficient is not statistically significant for any shooters when using w=1 (dark blue). For shooters with detectable fluctuations in performance, signal-to-noise ratio for performance prediction coefficient generally increases with increasing window size.

(c) Across subject mean estimates of performance prediction coefficient indicate strong presence of performance fluctuation among subjects. Performance prediction coefficients averaged across participants indicate that performance fluctuation is a strongly prevalent phenomenon among shooters. Increasing estimates of performance prediction coefficient with increasing window sizes, and approximately 2x increase in performance prediction coefficient for continuous shot outcome compared to binary shot outcome are both consistent with relationships predicted from simulation. Mean estimates of continuous shot outcome performance prediction coefficients at w=10 and w=20 are significant at p<1e-6 and p<1e-5 respectively. Mean estimates of performance prediction coefficients on discrete shot outcomes are all significant at p<0.01. Error bars represent 95% conf. int.

(d) Estimates of performance prediction coefficients individuals shows strong evidence of widespread performance fluctuation among individuals. 11 out of 16 participants showed strong evidence of fluctuations when calculating performance prediction coefficient in their sequences of continuous shot outcomes. We also find strong evidence of performance fluctuation in combined shot data across shooters when controlling for individual mean differences in shooting ability before combining their data, also indicating that fluctuating performance is widespread among shooters. Error bars represent 95% conf. int. calculated by bootstrap with n=10,000.

(e) Summary scatter plot of statistical significance of performance prediction coefficient measurement in individual shooters. Note that both horizontal and vertical axes are double-log scaled to aid visualization, and gray lines indicate thresholds for significances for both axes. Data suggest incredibly strong evidence that a majority of tested shooters display performance fluctuation, especially for continuous shot endpoint based estimates of performance prediction coefficient using w=10 (green) and w=20 (yellow). For reference, the 3 participants displaying the strongest evidence of performance fluctuation for w=20 had p<1e-23. The majority of continuous shot outcomes estimates are far above the unity line (black) indicating vast improvement in statistical strength compared to discrete shot outcomes.

(f) Estimated excursion size in units of make percentage for individual shooters. We estimated excursion size from individual participants' performance prediction coefficients estimated on continuous shot outcome with w=10 using simulated relationships between performance prediction coefficient and excursion size. Average excursion size across participants and estimated effect size for combining sequences across shooters are much larger than previous work. Error bars indicate 95% conf. int. obtained from bootstrap with n=10,000.

Previous work which has struggled to detect streakiness in individuals has also estimated that average excursion sizes for performance streakiness is relatively small, with estimates usually <5% (Arkes, 2010, 2013; Lantis & Nesson, 2021; Pelechrinis & Winston, 2022; Yaari & Eisenmann, 2011). Given the strong detection of streakiness in individuals, we next estimated the excursion size for individual shooters. To do this we interpolated a detailed simulation of the relationship between excursion size and performance prediction coefficient for continuous shot outcomes and w=20 (see Fig. 3.2d). In this relationship performance prediction coefficients is positively correlated with excursion size. Across participants, average estimated excursion size was 8.3%±1.9% [mean±95% CI], though estimates of excursion size for individual shooters covered a wide range. The streakiest shooter in our dataset had an estimated excursion size of 14.6%± 2.7%, while the shooter with the smallest detectable fluctuations had an estimated excursion size of 6.1%±2.8%. Even the least streaky shooter in our dataset displays excursions that are larger than those estimated in previous work. For reference, among players with at least 150 free-throw attempts in the 2020-2021 season, the difference between the average free-throw percentage (80%) and the best free-throw shooter that season (Chris Paul, 93%) was 13% (https://www.basketball-reference.com), indicating that the estimated range of excursion sizes for shooters in our dataset (6.1% to 14.6%) are highly meaningful in the context of actual basketball games.

3.3.4 Timescales of streaky performance

Seeing the relatively large excursions estimates for individual shooters, we wondered how quickly or slowly these shooters might switch between hotter and colder states of ability level. To better understand the timescales for detected performance fluctuations, we calculated predictable variability in continuous shot outcome across blocks and days in individual shooters by regressing average performance in single blocks or days onto average performance in the previous block or day. We also calculated shot to shot predictable variability during estimation of performance prediction coefficients for individual shooters. We converted R^2 from these regressions (adjusted for low n in day to day regressions) to units of cm^2 using the estimated variance at each timescale. Fig. 3.4a shows total predictable variability across these

three timescales, which indicates the contribution of performance fluctuations to individual shooter variability. Shooters with larger total contributions of streakiness to their variability tend to have larger contributions at each timescale (Fig. 3.4a), though averaged across participants there is not a detectable difference in the overall makeup of variability from streakiness (Fig. 3.4b). Interestingly, shooters also vary in the fractional contribution of streakiness at different timescales (Fig. 3.4c). Shooters in Fig. 3.4 a & c are sorted by total contribution of streakiness to variability, which highlights a pattern where the fractional contribution of shot to shot streakiness to variability decreases as total predictable variability increases. This suggests that higher levels of total predictable variability from streakiness among individual shooters are likely driven by increases in longer timescale streakiness more so than shot to shot streakiness.

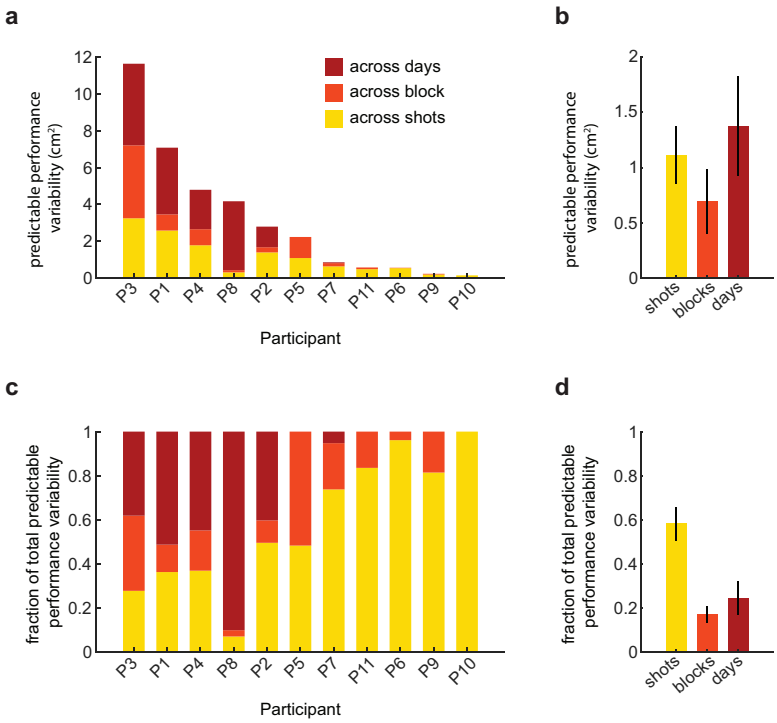

Figure 3.4: Streakiness varies across participants in its contribution to performance variability at different timescales

To investigate the timescales of performance streakiness we calculated how much from shot to shot (yellow), block to block (bright red), and day to day (dark red) variability in performance can be explained by the relationship between past performance and future performance for participants with detected individual streakiness.

(a) Total predictable variability across different timescales in individual shooters. Shooters vary widely in the total contribution of performance streakiness at different timescales to their overall variability. Shooters with greater total predictable performance variability generally show higher levels of predictable variability at all three timescales. Estimates of predictable variability were adjusted values to account for low number of data points when analyzing block to block and day to day data.

(b) Mean predictable performance variability across subjects. Data show no detectable differences between the average contributions of streakiness at different timescales to performance variability.

(c) Fractions of total variability accounted for by performance streakiness at different timescales. Individual participants also differ in the fractional makeup of their total predictable performance variability. Shot to shot streakiness tends to dominate total streakiness-driven variability for shooters with low overall predictable performance variability as seen in **a**, while the contribution of day to day streakiness is markedly increased for shooters with larger total predictable performance variability in **a**, suggesting that higher levels of streakiness among individuals is driven more by long timescale streakiness.

(d) Across participant mean fraction of total variability accounted for by streakiness at different timescales. For an average streaky shooter, shorter timescale streakiness is the major contributor to overall streakiness driven variability.

3.4 Discussion

We introduce a new method for detecting streakiness in shooters which calculates shooters' performance states as average performance over a window of shots, and calculates a performance prediction coefficient as the slope of the relationship between recent performance state and performance on the following shot. We combine this method with continuous-valued shot outcome to vastly improve statistical power for detecting streakiness at the individual level. We provide the first definitive evidence of streakiness in individual shooters with strong statistical confidence. Further, we find streakiness to be a widespread phenomenon among shooters, as we detect it in 11 of 16 shooters who partook in our study.

3.5 Methods

3.5.1 Stereo-vision tracking of ball approach and shot outcomes

A pair of well-calibrated stereo-vision cameras captured synchronous images of the ball at 80 frames/sec as it approached the rim. Ball position was tracked on individual images and then combined across cameras to produce 3D estimates of ball position with an estimated precision of 80 μm. Ball positions measured just before rim impact, or visual overlap with the rim in the camera's view, were combined with an estimate of average ball acceleration to project out the ball's trajectory and calculate shot endpoint an estimated precision < 0.5 mm. Camera images were also used to determine whether each shot was a make or miss based on ball position relative to the rim.

Categorizing made shots as direct or indirect makes

Each made shot was labelled as either a direct make or indirect make. Indirect makes are lower accuracy shots which go in after multiple bounces on the rim, while direct makes are high accuracy shots which go through the rim with minimal rim contact. Direct makes were identified as shots which pass through the

rim within 150 ms of their arrival time at the rim, which show downward post-impact velocities if impact with the rim occurred. Made shots not falling within this criteria were labelled indirect makes.

Fitting of classification model parameters

We designed a classification model which separates direct makes from other shots based on shot endpoint and approach angle of the incoming ball. This model implements an elliptical boundary where only shots with endpoints falling inside the boundary are labelled direct makes. We include parameters allowing modification of the ellipse geometry based on the approach angle of the ball on each shot. These parameters were linear functions of approach angle and constituted lateral and longitudinal shifts in the center position of the ellipse and gains on the lateral and longitudinal ellipse radii. A custom error function was defined with continuous valued errors for misclassifications occurring within α of the make region boundary and constant valued errors otherwise. This allowed use of gradient descent methods for estimating parameter values.

3.5.2 Shot data collection

We recruited 16 university students with previous organized basketball experience, of which 12 were members of club basketball teams that engaged in shooting practice. We recorded 3,200 free-throw shots shots from each individual shooter over the course of six sessions that always occurred within a span of two weeks while only allowing one shooting session per day. Shooting sessions consisted of shooting 500 free throws (700 on day 1) in short blocks of 50 shots separated by 2 min rest breaks, when shooters were given freedom to sit, stretch, use the bathroom, hydrate, and do whatever they needed to rest. None of the shooters reported overly tired at the end of sessions, and a brief analysis of changes in shooting rate throughout sessions and make percentage throughout blocks and sessions did not indicate any fatigue effects.

Free-throw shots were collected in a large experiment room with sufficiently high ceilings. Shots were of regulation dimension, taken from a distance of 15 ft from the backboard with a rim at a 10 ft height from

the ground. Participants were shown a running count of their makes and misses after each shot. This data was originally collected as part of a different study looking at the effects of showing shooter's a visual map of their accuracy after each shot. Analysis of shot data showed no significant changes in average shooter performance between the first and last sessions, indicating minimal effect of this visual feedback.

3.5.3 Simulating streaky shot sequences

We used two types of patterns to model changes in underlying ability which manifest in sequences of shots. For random jumps we assumed a constant transition probability p for switching from the hot state to the cold state and vice versa. For the random walk we simulated a random walk with a tether coefficient B which continuously pulls the random walk back towards zero. To produce patterns of a specific excursion size, simulation patterns were scaled to match the RMS value of a sine wave of amplitude equal to the desired excursion size. As both the random jump and tethered random walk pattern have exponential autocorrelation functions, we matched the timescales across both patterns by equating their autocorrelation functions for a chosen average periodicity. Simulations were done using an average period of 400 shots with mean shooting percentages set to 70%.

For discrete shot outcomes, these patterns were taken to directly be underlying "true" make percentages describing the shooters performance state, and we simulated random binomial trials with success probability on each trial equal to the value of the pattern make percentage on each shot. For continuous shot outcome, we drew the simulated shot endpoint error on each shot by drawing for a random normal distribution with sigma inversely correlated with the expected true make percentage indicated by the pattern.

3.5.4 Aggregating data across shooters

Before aggregating data across individuals, we normalized individual shooter shot outcomes by their average performance. For binary shot outcomes, we subtracted the average direct make fraction from each binary sequence of direct makes and misses.

3.6 References

Arkes, J. (2010). Revisiting the Hot Hand Theory with Free Throw Data in a Multivariate Framework. *Journal of Quantitative Analysis in Sports*, *6*(1). https://doi.org/10.2202/1559-0410.1198

Arkes, J. (2013). Misses in "Hot Hand" Research. *Journal of Sports Economics*, *14*(4), 401–410. https://doi.org/10.1177/1527002513496013

Avugos, S., Bar-Eli, M., Ritov, I., & Sher, E. (2013). The elusive reality of efficacy-performance cycles in basketball shooting: An analysis of players' performance under invariant conditions. *International Journal of Sport and Exercise Psychology*, *11*(2), 184–202. https://doi.org/10.1080/1612197X.2013.773661

Bar-Eli, M., Avugos, S., & Raab, M. (2006). Twenty years of "hot hand" research: Review and critique. In *Psychology of Sport and Exercise* (Vol. 7, Issue 6, pp. 525–553). https://doi.org/10.1016/j.psychsport.2006.03.001

Chaisanguanthum, K. S., Shen, H. H., & Sabes, P. N. (2014). Motor variability arises from a slow random walk in neural state. *Journal of Neuroscience*, *34*(36), 12071–12080. https://doi.org/10.1523/JNEUROSCI.3001-13.2014

Cohen, R. G., & Sternad, D. (2009). Variability in motor learning: Relocating, channeling and reducing noise. *Experimental Brain Research*, *193*(1), 69–83. https://doi.org/10.1007/s00221-008-1596-1

Gilovich, T., Vallone, R., & Tversky, A. (1985). The Hot Hand in Basketball: On the Misperception of Random Sequences. In *COGNITIVE PSYCHOLOGY* (Vol. 17).

Haberstroh, T. (2017, June 7). He's heating up, he's on fire! Klay Thompson and the truth about the hot hand. *ESPN*. https://www.espn.com/nba/story/_/page/presents-19573519/heating-fire-klay-thompson-truth-hot-hand-nba

Jagacinski, R. J., Newell, K. M., & Isaac, P. D. (1979). Predicting the Success of a Basketball Shot at Various Stages of Execution. *Journal of Sport Psychology*, *1*, 301–310.

Klay Thompson sets NBA record with 37 points in a quarter. (2015). *Bleacher Report*. https://bleacherreport.com/articles/2341021

Kojima, S., & Doupe, A. J. (2011). Social performance reveals unexpected vocal competency in young songbirds. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1687–1692. https://doi.org/10.1073/pnas.1010502108

Lantis, R., & Nesson, E. (2021). Hot Shots: An Analysis of the "Hot Hand" in NBA Field Goal and Free Throw Shooting. *Journal of Sports Economics*, *22*(6), 639–677. https://doi.org/10.1177/15270025211006889

Miller, J. B., & Sanjurjo, A. (2014). A Cold Shower for the Hot Hand Fallacy: Robust Evidence that Belief in the Hot Hand is Justified. *IGIER Working Paper #518*. https://ssrn.com/abstract=2450479

Miller, J. B., & Sanjurjo, A. (2018). Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers. *Econometrica*, *86*(6), 2019–2047.

Miller, J. B., & Sanjurjo, A. (2019). *A Cold Shower for the Hot Hand Fallacy: Robust Evidence that Belief in the Hot Hand is Justified*.

Miller, J. B., & Sanjurjo, A. (2021). Is it a fallacy to believe in the hot hand in the NBA three-point contest? *European Economic Review*, *138*. https://doi.org/10.1016/j.euroecorev.2021.103771

Morgulev, E., Azar, O. H., & Bar-Eli, M. (2020). Searching for momentum in NBA triplets of free throws. *Journal of Sports Sciences*, *38*(4), 390–398. https://doi.org/10.1080/02640414.2019.1702776

Pelechrinis, K., & Winston, W. (2022). The hot hand in the wild. *PLOS ONE*, *17*(1), e0261890. https://doi.org/10.1371/journal.pone.0261890

Sternad, D. (2018). It's not (only) the mean that matters: variability, noise and exploration in skill learning. In *Current Opinion in Behavioral Sciences* (Vol. 20, pp. 183–195). Elsevier Ltd. https://doi.org/10.1016/j.cobeha.2018.01.004

Wardrop, R. L. (1999). *Statistical Tests for the Hot-Hand in Basketball in a Controlled Setting*.

Yaari, G., & Eisenmann, S. (2011). The hot (invisible?) hand: Can time sequence patterns of success/failure in sports be modeled as repeated random independent trials? *PLoS ONE*, *6*(10). https://doi.org/10.1371/journal.pone.0024532

Appendix A: Supplementary materials for Chapter 1

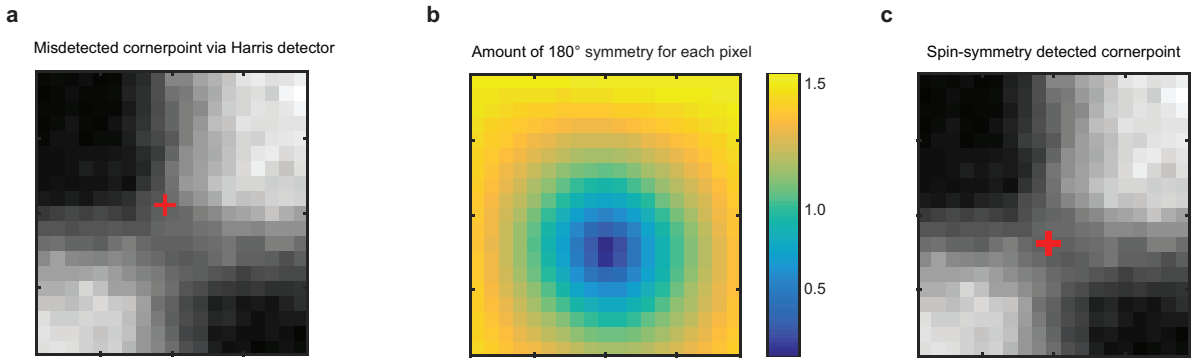

Figure A1: Improved chessboard cornerpoint detection accuracy by leveraging local 180º rotational symmetry in images affected by blur

To monitor the basketball's motion after release across multiple captured frames and for shooters with different release heights, the cameras in our system are placed 10-20 ft from where the ball is typically released. This increases the trackable volume of our system but requires the cameras to be focused at further distances, causing the close-range images required for calibration to be blurry. As traditional cornerpoint detections were sensitive to blur, we developed a new method, which we refer to as the "spin-symmetry" detector, which defines the cornerpoint location to be the point of maximum local 180º rotational symmetry.

(a) Harris cornerpoint detection algorithm misdetects the cornerpoint location in an image patch containing a chessboard cornerpoint in a somewhat blurry calibration image. Traditional computer vision toolboxes use the Harris corner detection algorithm, which uses the eigenvalues of the Hessian matrix to identify cornerpoints as locations of high change in brightness in both horizontal and vertical directions along the image. However, in blurry images, such as the image patch around a cornerpoint shown in here, the true cornerpoint is obscured by a "flat" gray region of more uniform brightness. In this case, the Harris detector misdetects the cornerpoint to be at the edge of the blurry region, which can often be a detection error of several pixels.

(b) Local 180º rotational asymmetry in the image around individual pixels. Instead of using the Harris detector, we leveraged the local rotational symmetry of cornerpoints in the image. We calculated the 180º rotational symmetry of the local image surrounding the cornerpoint within the image as the sum-of-squared-differences between the rotated and unrotated image patch when rotated around a given pixel. This allows better localization of the cornerpoint in blurry images. To provide subpixel accuracy in localizing cornerpoints, we calculated the local symmetry for subpixel shifts of the centerpoint using 2D-interpolation, which was used in a gradient descent algorithm to compute the optimal centerpoint with lowest rotational asymmetry.

(c) Improved cornerpoint localization when using spin-symmetry detector in example image. Using the spin-symmetry detector, the detected cornerpoint (red) is no longer at the edge of the "flat" grey region due to the image blur, and is now better situated at the intersection of the chessboard squares.

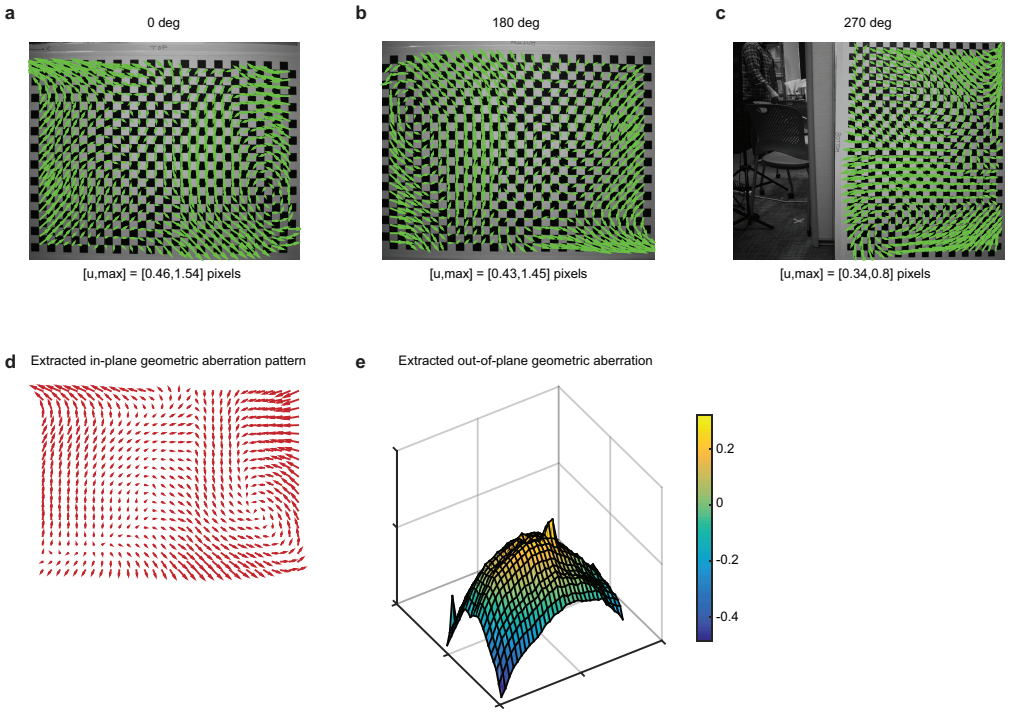

Figure A2: Refined detection of chessboard grid geometry to account for physical warping and imperfectly printed grid pattern

Standard camera calibration procedures use pictures of flat-mounted chessboard grids to determine camera-to-camera positioning and orientation, and to measure image distortion due to warping from wide-angle camera lenses. This procedure assumes that the detected points on images of the chessboard grid come from a grid with perfectly regular spacing of the intersection points. However, despite careful mounting of chessboard grids on flat table-top glass, and despite printing the chessboards with high-resolution printers on polystyrene sheets, we noted that there were large, smoothly patterned calibration errors that were consistent in their direction relative to the chessboard geometry across multiple images.

(a-c) Images of chessboard grids used in calibration of lens distortion parameters with overlaid reprojection errors. Standard calibration procedures use the discrepancy between model parameter predicted chessboard cornerpoint locations, assuming a perfect grid geometry, and detected cornerpoint locations on the images to drive optimization of model parameters. The discrepancy between parameter predicted and image detected cornerpoint locations are referred to as reprojection errors. On images of chessboard grids used in calibrating lens distortion parameters, we saw smooth patterns in reprojection errors (green) that were consistent with the geometry of the chessboard grid, rather than the image coordinates themselves, evidenced by the rotation of the smooth pattern of reprojection errors consistent with rotation of the calibration rig across images.

(d-e) Extracted and refined calibration rig reference geometry. We implemented a gradient descent algorithm which iteratively changed the 3D position of each cornerpoint in the calibration rig reference geometry, both in-plane and out-of-plane, by minimizing the reprojection errors when fitting the calibration parameters.

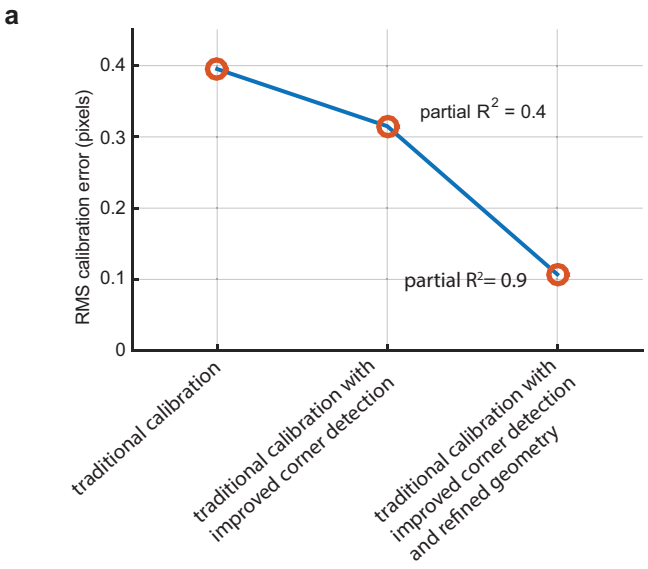

Figure A3: Summarized improvements to standard camera calibration procedure

Overall, use of the spin-symmetry detector and extraction of the geometry refinement when fitting both the camera lens distortion parameters and the camera-to-camera geometry reduced the errors in calibration by ~90% compared to standard methods.