ASSESSING THE RELIABILITY AND VALIDITY OF ONLINE TASKS TO ASSESS PERCEPTUAL COGNITIVE SKILLS IN BASEBALL

by

Georgia Grieve

B.PubHealth., The University of Canberra, 2016

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(KINESIOLOGY)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2022

© Georgia Grieve, 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis entitled:

Assessing the reliability and validity of online tasks to assess perceptual cognitive skills in baseball

submitted by Georgia Grieve in partial fulfillment of the requirements for

the degree of Master of Science

in Kinesiology

Examining Committee:

Dr Nicola Hodges, Professor, School of Kinesiology, UBC
Supervisor

Dr Romeo Chua, Professor, School of Kinesiology, UBC
Supervisory Committee Member

Dr Miriam Spering, Associate Professor, Faculty of Medicine, UBC
Supervisory Committee Member

Additional Supervisory Committee Members:

Dr Sean Müller, Associate Professor, School of Science, Psychology and Sport, Federation University
Supervisory Committee Member

Abstract

There has been an increasing interest in training perceptual skills in sports through video-based methods, particularly in baseball (e.g., baseball pitch recognition). However, there is little empirical evidence related to their reliability and validity, to help guide the efficacy and application of video-based training methods. Here we sought to investigate whether an online task can be used to a) assess perceptual cognitive skills and b) discriminate across age and different skill groups. An online experimental platform was developed to help collect data on baseball specific predictions concerning pitch type and dynamic visual acuity among baseball players (n = 21) and novices (n=14) with low baseball experience (age 13 - 22 yr). The baseball prediction task consisted of two different pitchers, three pitch types and three occlusion points (such that ball flight information was progressively removed). Each clip was shown twice. The prediction task was shown to be valid and reliable, discriminating across skill groups and showing reliability across repeated viewings and from the online task to an in-person assessment (n=6). There were, however, differences in reliability and discriminant validity based on the type of pitcher, with one pitcher being responded to more accurately and reliably. The skilled participants showed good discriminability between fastballs and change-ups versus the novice group and there was a trend for better dynamic visual acuity in the skilled group. Due to difficulties in recruitment, the data was mostly based on adult participants, such that further data collection is needed to make conclusions about age-group differences and development of perceptual-cognitive skills in baseball. The current data mostly help to provide baseline data about current stimuli and methods for assessment of perceptual skills in baseball, showing evidence speaking to the validity and reliability of online methods.

Lay Summary

There has been an increasing interest in the commercial training of perceptual-cognitive skills in sports through video-based methods, particularly in baseball (e.g., pitch recognition). However, there is little empirical evidence as to the effectiveness of online training apps, nor research into pitch recognition skills in players across different ages and experiences. We launched an online study to assess pitch anticipation skills and dynamic visual acuity from baseball players of different ages and experiences. The baseball prediction task was shown to be valid and reliable, discriminating across skill groups. Due to difficulties in recruitment, the data was mostly based on adult participants, such that the current data help to inform methods for continued study in this field. Our aim is to develop reliable and valid tests which can be used to study how perceptual-cognitive skills develop in baseball and ultimately, the effectiveness of online tasks for training and transfer to baseball.

Preface

This thesis is an original intellectual product of the author, G. Grieve. All work presented was conducted with support from personnel in the Motor Skills Laboratory in the School of Kinesiology at the University of British Columbia. The project and methods reported were approved by the University of British Columbia's Behavioural Research Ethics Board [certificates H19-02705; online study; H21-03828, in-person testing].

I was the lead investigator, and was responsible for stimuli collection, data collection and analysis, and writing. Z. Besler was responsible for the creation of the online task delivered via Gorilla and collecting data for the in-person portion of data collection. N.J. Hodges was the supervisory author on this project and was involved in concept formation, data analysis and writing.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my supervisor, Dr Nicola Hodges, for her continuous support, guidance, and patience on this project through the many changes that were made as challenges arose. I am extremely grateful for the opportunities I have been afforded, and for the time spent learning from and working with Dr Hodges.

Thanks also to my committee members, Dr Miriam Spering, Dr Romeo Chua, and Dr Sean Müller, who shared their knowledge and experience to help guide this project.

The Motor Skills Lab at UBC is full of wonderful people who were of great help throughout my Master's. A special thanks to Zachary Besler, who developed the online task for this study and has generally contributed greatly to this study.

A great thanks is owed to my parents, brother, and friends who stuck through this and encouraged me all the way.

Dedicated to Leanne and Paul. Patience grasshopper.

Chapter 1: Introduction

1.1 Perceptual-cognitive skills in sport

Perceptual-cognitive skills have had a long history of study in sports and have been defined as skills that guide the use of environmental information, underpinning action anticipation, selection and execution (see Hodges et al., 2021; Williams et al., 2011, for reviews). They have sometimes been defined as abilities, rather than skills, as there is a continuing debate as to how much these skills are dependent on experience versus more innate ability. This is particularly true for visual “skills”, such as static and dynamic visual acuity (Hodges et al., 2021). Because these primarily visual perceptual skills underpin action-related decisions, these skills have been referred to in the literature as perceptual-“cognitive”. Moreover, as the assessment of these skills can be made in the absence of action execution, they have been distinguished from motor skills. This doe not mean that these skills are independent of action, nor that the cortical motor system is not involved in their functioning, just that they can be considered separately from the technical enactment. In baseball, perceptual-cognitive skills have primarily been studied with respect to three types of information use: situational or contextual cues (e.g. pitch count; Gray & Canal-Bruland, 2018), body-related kinematic cues (e.g. wrist position indicating pitch type, Muller et al., 2017), and object tracking (e.g., visual gaze of ball flight; Fooken & Spering, 2020). The ability to use these sources of information has an impact on the success of a player at-bat (e.g., Muller & Fadde, 2016).

In 2018, Joey Votto led the MLB (Major League Baseball) in on-base percentage and runs created (31% higher than the MLB average), had the lowest O-swing % (i.e., how often he tried to hit a ball that was going outside the strike zone), a 5:1 Z-swing to O-swing ratio (i.e., how often he swung at a ball in the strike zone compared with outside of the strike zone), and a well above league average 19% walk rate (i.e., how often he walked to first base as a result of the pitcher throwing four foul balls outside the strike zone) (Adler, 2019; Dokken, 2019). Although this was one of Votto’s least successful seasons, these statistics reflect a baseball player who is either a very good guesser, or much more likely, a batter who has highly developed perceptual-cognitive skills (Adler, 2019; Dokken, 2019). Votto’s statistics suggest that he is able to use these skills to be aggressive in the zone when the pitch is right. An example player like Votto leads us to ask how perceptual cognitive skills can be trained to make such season performances consistent and perhaps offer hitters an opportunity to keep up with advances on the pitching side of the game.

A hitter who relies just on ball flight to respond, will have limited time available to execute their response to a pitch. Research in striking sports has shown that experts and near experts are able to make an accurate prediction about delivery type before ball release or very early in ball flight and can then make a decision about the location of this delivery shortly after ball release (Bahill & Karnavas, 1993; Fadde, 2006; Gray, 2010; Morris-Binelli et al., 2018). Thus, the earlier a batter can identify the pitch type and direction, the earlier they can initiate their swing, or decide not to swing. It appears that Joey Votto was able to accurately anticipate pitches at a level higher than other batters, so that he could leave those pitches which would be outside the zone and be aggressive with those pitches landing in his preferred strike zone. Research in various interception sports has shown that experts are able to use advanced cues from the body before ball release, and in the early stages of ball flight to guide their response, setting elite athletes apart from near experts and novices (e.g., Abernethy, 1996; Fadde, 2006; Gray, 2010).

There has been a relatively recent interest in using virtual reality, virtual environments, and video game-type tools to train perceptual skills in athletes (for example, Gamesense; https://gamesensesports.com/for-players/). However, the research is lagging behind the practice, so there are no evidence-based guidelines regarding when and how to best introduce this training and limited evidence to show far transfer from playing video-type games to performing in competitive on-field situations (Fadde, 2006; Gray 2017; Morris-Binelli & Muller, 2017; for a review, see Gray, 2019). In order to better determine whether there are specific ages when perceptual skills' training is best introduced, it would be beneficial to know how these skills typically develop in youth athletes before intervening.

1.1.1 Temporal occlusion to assess anticipation

In order to assess participants' ability to use advance cues to accurately anticipate, temporal occlusion methods have been employed (Jones & Miles, 1978; for recent reviews see Muller & Abernethy, 2012; Smeeton et al., 2019 in Williams & Jackson). A given action is filmed, for example a baseball pitch and the footage is edited at particular points to withhold or provide information deemed important to anticipation. An edited video is then shown to an athlete and accuracy in responding is determined as a function of the amount and/or type of information available, often through skill-group comparisons. An improvement in prediction accuracy between occlusion points is considered to be evidence of information becoming available during the given viewing window (e.g., Farrow et al., 2005). Prediction results above chance (e.g. above 33% in a three-outcome task) can be used to establish that there is information available in a given window of the action to assist in determining the action's outcome (Farrow et al., 2005). Studies using temporal occlusion to assess anticipation show that experts are consistently better able to anticipate outcomes than novices and that they are better able to make use of advanced kinematic information from the body of an opponent, which go undetected by novices (e.g., Abernethy & Russell, 1987a, 1987b; Farrow et al., 2005; Aglioti et al., 2008; Farrow & Reid, 2012; Müller et al., 2017).

The temporal occlusion method as a way of measuring differences across skill groups, inferring what information is used for anticipation and even training athletes, has come under some criticism (e.g., Davids et al., 2013). Farrow et al. (2005) discussed two potential issues associated with the temporal occlusion method. The first pertained to a potential confound related to the duration of the viewing period and whether an increased duration leads to better accuracy because more information is available to the viewer or because the viewer has repeatedly seen the earlier information in other edited clips. However, when the typical progressive temporal occlusion was compared to a window-based paradigm (controlling for exposure to prior information), there was no significant difference in accuracy at any given time point (Farrow et al., 2005). The second potential issue with temporal occlusion has to do with a lack of ecological validity (Davids et al., 2013). Temporal occlusion is often used in a laboratory setting, aiding experimental control, but removed from its natural setting where response time constraints are real and information is continually available. Further, a verbal or written response is often required, rather than a perceptually-coupled response to the action (e.g. swing a bat in an attempt to hit a pitch). There is evidence that the expert advantage is enhanced in the coupled setting and as such, requiring an uncoupled response may weaken evidence of differences (Farrow et al., 2005; Mann et al., 2010). However, through the use of liquid crystal spectacles in a real tennis serve return task, Farrow and colleagues (2005) showed no significant difference between predictions made in response to a real serve and those to video serves. As such, this video occlusion method has generally been thought of as an appropriate method to study perceptual-cognitive skills of athletes, but limits to this method need to be considered.

1.1.2 Expert-novice differences in prediction accuracy

In relatively early studies using the temporal occlusion method, significant differences between anticipatory abilities of experts and novices were demonstrated in badminton (Abernethy & Russell, 1987a, 1987b). Experts were able to make accurate predictions earlier in the action than novices, indicating the use of body-related kinematic cues well before contact with the shuttle was made (Abernethy & Russell, 1987a). Further, experts were able to more effectively reduce their prediction errors as more information became available, while novices were only able to do this closer to contact being made (Abernethy & Russell, 1987a). These differences between skilled and less skilled athletes in prediction accuracy have been replicated in many other sports and athlete groups; particularly in sports requiring interceptive actions; including basketball (Aglioti et al., 2008), cricket (Brenton et al., 2016; Mann et al., 2010; Müller et al., 2006; Weissensetiner et al., 2008), baseball (Canal-Bruland et al., 2012;

Müller et al., 2017; Muraskin et al., 2015; Ranganathan & Carlton, 2007), and squash (Abernethy et al., 2001).

The temporal occlusion method is not always used in isolation to infer what information is used to make predictions. Another measure which has been relatively common in sport-anticipation research is eye tracking (see Gegenfurtner et al., 2011; Hodges et al., 2021 for reviews). Eye tracking is thought to give more refined information about how accuracy in predictions are achieved, through tracking of gaze throughout an action, such as determining where an expert is looking on the body of a server in tennis before the serve is made (Ward et al., 2002). Although there have been a number of studies showing differences across skill groups in terms of where they are looking (e.g., Kato & Fukuda, 2002; Savelsbergh et al., 2002; Takeuchi & Inomata, 2009; Williams & Davids, 1998), there are also studies failing to show differences in terms of fixations (both location and duration; e.g., Abernethy & Russell, 1987b; Wood & Abernethy, 1997). It may be that novices know where to look but are unable to use the information to effectively guide anticipation. Fixations by themselves also fail to show information that is picked up in the periphery, whereby an elite athlete may be merely anchoring their gaze to maximize the amount of information which can be processed (Vater et al., 2020).

In striking sports, there is evidence that highly skilled, near to expert players, are significantly better at predicting the outcome of an opponent's action before ball release in bowling, serving, or pitching actions (e.g., Brenton et al., 2016; Farrow & Reid, 2012; Müller et al., 2006; Weissensteiner et al., 2008). Highly skilled players were able to predict outcomes above chance at the point of ball release, while novices were only able to predict above chance when there was no occlusion of the bowler's action. These data again indicate that lesser skilled players and novices are relying on ball flight to make predictions, rather than body-related kinematic cues from the pitcher/thrower (Brenton et al., 2016; Müller et al., 2006).

Not only are more skilled athletes able to use advanced kinematic information to guide action selection, but they are able to use it to guide an inhibitory strategy to correct their chosen swing pattern (Gray, 2010; see also Nakamoto et al., 2013). Gray (2010) showed that near-expert batters were using ball flight information differently than lesser skilled batters; adjusting their swing to produce a more advantageous launch angle from a less than ideal pitch rather than guiding the decision whether to swing or to stop a swing. When assessing visual anticipation in professional baseball batters, Morris-Binelli and colleagues (2018) showed that while accurate predictions could be made about pitch type early in the action, more time and ball flight was required to accurately predict the final location of the ball (and make important adjustments). Batters completed a task requiring them to respond about both pitch type and pitch location when shown occluded footage of two minor league pitchers. At the first three occlusion points (ball release, ball release + 80 ms, ball release + 200 ms), pitch type predictions were more accurate than pitch location predictions and this did not change until all ball flight information became available to the batter (no occlusion condition). Further, paired predictions (location and type) were less accurate at all occlusion points than location or type individually.

Asking batters to decide location before they have determined pitch type might interfere with the batter's ability to accurately predict the ball's flight. Because pitch location accuracy was not above guessing level until 200ms after ball release in the study by Morris-Binelli and colleagues (2018), it appears that more ball flight is required for a decision to be made about the final location of the ball relative to the strike zone and that batters are likely to be deciding about pitch type first. Further, the professional batters in Morris-Binelli and colleagues' (2018) study were able to accurately predict the pitch type at ball release. This decision hierarchy is also in line with Bahill and Baldwin's (2004) discussion that the pitch can be broken down in to three parts: the collection of sensory information, followed by the processing of when and where contact will be made, and finally, the swing. This order of events suggests that pitch type needs to be discerned first in order to determine where and when contact will potentially occur, and thereby whether contact should in fact occur (see also Bahill & Karnavas, 1993). Correct predictions 200 ms after ball release were also found to be negatively correlated with strikeout rate in games and positively correlated with walk-to-strikeout ratio and on base percentage (Morris-Binelli et al., 2018). The accurate prediction of pitch type at 80 ms after ball release was positively correlated with slugging percentage, a statistic related to power hitting. Based on these relationships, it may be expected that early, accurate anticipation of pitch type is the crux of hitters being able to be aggressive in the zone (Morris-Binelli et al., 2018).

1.1.3 Differentiating the role of action and perceptual experiences in action prediction

Research comparing highly experienced athletes to less experienced athletes or novices, has shown that motor (i.e., physical) experience improves the ability to predict the outcome of a given action from temporally occluded videos (e.g., Aglioti et al.,2008; Urgesi et al., 2012; Mulligan et al., 2016a). When basketball players were compared to experienced watchers (journalists) and novices, they were better able to predict the outcome of a free throw at various occlusion points, particularly in advance of ball release (Aglioti et al., 2008). Novices were found to be more likely to respond 'I don't know' to ball outcome success questions, before ball release, relying on ball trajectory to guide their predictions. This increased accuracy and certainty in the motor experts compared to the novices, indicated that the motor experience of the athletes allowed them to use early kinematic information effectively to guide their decisions. Aglioti and colleagues provided support for this motor-based interpretation via a TMS

(Transcranial Magnetic Stimulation) experiment. The muscles involved in the free-throw shot, showed high levels of activation in the athletes (activation via TMS and assessed via EMG) compared to novices and also in comparison with a control condition which involved watching a soccer kick. Although the muscle activations of the athletes were not different to expert watchers, only for the athletes were activations directly related to the success of the shot (Aglioti et al., 2008).

There is evidence that if, during action observation, the muscles involved in the observed action are activated in a manner incongruent to the action being watched, an experienced person's ability to predict the outcome of the action is negatively impacted (Mulligan et al., 2016a). This interference is not seen in an attentional-control tone monitoring task, or during mimicry of the observed action. This study and others (e.g., Mulligan et al., 2016b), suggest that the motor recruitment required by the secondary force production task interfered with the activation of the person's motor system, which in turn hindered the perceptual prediction skills of the experienced group (Mulligan et al., 2016a). These results are in line with what is suggested by common coding theory; that the experience of performing an action is bidirectionally linked to the ability to anticipate the outcome of the action (Hommel et al., 2001; Prinz, 1997). In terms of how this is possible, there is neurophysiological evidence that individuals with motor experiences activate the same areas of the brain when watching others perform actions as those that are activated when they themselves perform the same actions (e.g., Calvo-Merino et al., 2005). This activation of motor areas during observation is known as motor simulation or resonance and is based on evidence of what has been termed a mirror neuron system in the brain (Calvo-Merino et al., 2005; Rizzolatti & Craighero, 2004).

1.1.4 Development of perceptual-cognitive skills with age

There is limited investigation of the development of anticipation in striking sports with regards to age. One exception is a study by Farrow and Reid (2012), who provided evidence that adult experts were able to generate more sophisticated responses to an action, compared with youth "experts" (Farrow & Reid, 2012). Older tennis players (~ 17 years) were able to pick up on a service pattern, while younger players (~11 years), were not (Farrow & Reid, 2012). The first serve of every game was always the same and the older players were able to identify this pattern by the ninth service game and commence their response well before racquet-ball contact. The older group was also able to use pre-contact kinematic information to assist in their anticipation of first serve location, while the younger group was not. Farrow and Reid suggested that the differences between the age groups may be due to the fact that the skills typically required of the younger group when playing tennis do not require advanced perceptual skills. This could be due to slow service speeds or the unreliable or inconsistent kinematic information from opponents (Farrow & Reid, 2012). As discussed by French and McPherson (1999), athletes of a young age or playing at a low skill level are less likely to use advanced cues, because they simply do not need them. The ball does not travel as fast and deceptive tactics are not employed, so participants have enough time to coordinate their response.

One comprehensive developmental study of age and skill group differences in cognitive and visual-perceptual skills was conducted by Ward and Williams (2003). Youth elite and sub-elite soccer players aged 9-17 years were compared for static and dynamic visual acuity, depth sensitivity, peripheral awareness, anticipation, memory recall, and situational expectations. With respect to the non-sport specific skills; static visual acuity showed differences across age, irrespective of skill, especially comparing the 9 to the 13-year olds. However, skill group differences were shown at the U11 (under 11 years) and U13 ages. Peripheral awareness age-group differences were seen relatively early in the elite groups compared to the sub-elite groups. The sub-elite groups showed age-related differences later in development than the elite groups and by U15 and U17, there were no significant differences between skill groups. These data suggest that with age, peripheral awareness improves, but quality practice early can accelerate this acquisition.

With respect to age and skill differences in sport-specific skills, elite players showed significantly better anticipation skills when compared to sub-elite players at all ages (Ward & Williams, 2003). In younger age groups, elite players were better at recognising key players in a game situation than sub-elite. The elite players' ability to correctly identify key players remained consistent across all age groups, but the sub-elite group showed differences from U9 to U13, with accuracy getting better with increased age. Because this was a cross sectional comparison and practice history was not specifically measured, we can only speculate that the superior peripheral awareness and anticipatory skills of the more skilled players was due to earlier exposure to domain specific practice activities and that the uniform improvement in visual acuity (irrespective of skill), was unrelated to this practice.

In a study investigating the relationship between practice history and anticipatory ability in skilled and lesser skilled cricket players across three age groups (U15, U20, adult), Weissensteiner and colleagues (2008) reported that the older age groups had superior anticipatory skills compared to the U15 groups, regardless of skill level. The U15 groups did not demonstrate the use of advanced cues to guide anticipation and were unable to predict the outcome of a delivery above chance at ball release (when watching an adult swing bowler). The data collected about practice history showed that the U15 skilled group had spent up to three times as much time in organized cricket practice as the less skilled group, but this additional practice did not translate to perceptual-cognitive skill benefits, at least in terms of using advanced cues to guide anticipation. Although a strength of this study is that practice history was measured, the large age gap between the U15 and U20 age groups prevents conclusions about the age when the development of perceptual cognitive skills begins or is accelerated. Weissensteiner and colleagues acknowledged this limitation and recommended further research to understand exactly when perceptual cognitive expertise is developed and what in an athlete's practice history assists in the development of these skills. Moreover, they only measured skill development in anticipatory skill performance related to the sport but not other visual skills which could limit development and potentially explain age and skill group differences.

1.1.5 Development of visual skills

Hitters must respond to an object which is both low contrast and thrown from far away, while under significant time constraints (Laby et al., 2019). General visual skills, which are not specific to the sport, have been investigated in athletes and in the sport of baseball (see Hodges et al., 2021 for review). Laby and colleagues (2019) used a Landolt C task to assess dynamic visual acuity (DVA) in baseball players, in which the difficulty of the task was manipulated by changing the size, contrast, and viewing time of the target (i.e., the letter C which had different sized openings). Visual acuity was positively correlated with the baseball statistic of walk rate (and negatively correlated with swing rate), indicating greater selectiveness or discrimination at the plate when deciding whether to swing (Laby et al., 2019). Similar results were found by Burris and colleagues (2018) when they assessed the sensorimotor abilities of professional baseball players in comparison with their in-game statistics; sensorimotor skills were positively correlated to on-base percentage, walk rate, and strikeout rate.

Uchida and colleagues (2013) assessed whether baseball players had a superior ability to perceive moving objects when compared to non-players (i.e., DVA) using a Landolt C tracking task. The baseball players were better able to track the moving Landolt C rings compared to the non-player group. Because there was a lack of significant difference between the two groups in terms of retinal error, the authors concluded that the difference in tracking ability between groups was likely due to exposure to a specific stimulus (in this case, tracking a baseball), as opposed to an innate ability to move one's eyes well.

When testing baseball athletes' visual skills using the Nike Sensory Station battery, Klemish and colleagues (2018) showed that eye-hand coordination and go/no-go response time improved when comparing high school baseball athletes (as young as 14 years) to college-level and professional athletes. It is also worth noting that in this study their pitchers and hitters did not differ in their visual capabilities before the professional level (Klemish et al., 2018). While the study included high school players as young as 14 years, the purpose of the study was not to assess visual skills relative to age and skill, and as such, we are unable to draw any conclusions about the likely cause of these changes in the older, yet highly trained athletes.

In another study, baseball and non-baseball athletes' eye movement strategies were studied during the Landolt C DVA task (Palidis et al., 2017). DVA was linked to smooth pursuit kinematics and saccadic eye movements, with smooth pursuit tracking leading to worse accuracy at this task than a saccadic, anticipatory strategy. The more senior baseball players also made fewer reverse saccades than the junior players, leading the researchers to suggest that experience with baseball facilitates oculomotor choices and success on this DVA task (Palidis et al., 2017). What these data show is that training in baseball positively transfers to other tasks which assess the ability to track moving objects at fast speeds, and it is the eye movement strategies which appear to transfer.

1.2 Study aims and hypotheses

Based on the gaps in the literature discussed above, in my study I am seeking to do the following and address these aims through the study of developmental as compared to Varsity baseball athletes:

1) Compare sport-specific perceptual-cognitive (PC) skills related to pitch discrimination across athletes of different ages (U15 years, U17 years and over 19 years). In order to test the validity of the experimental task, I also intend to compare across skill (i.e., experienced/high performing versus novice performer). I expect there to be advantages for more experienced athletes versus younger and/or less experienced novice groups.

2) Evaluate age and skill group differences in non-domain specific general visual skills (i.e., dynamic visual acuity, DVA) as compared to domain-specific PC skills (i.e., baseball pitch discrimination) and analyse the relationship between these two skills. The evidence is quite mixed regarding experience-based differences in non-domain specific visual skills. If DVA is experience-dependent, advantages will be seen for the more experienced athletes versus younger and/or less experienced novice groups. I would also then expect that DVA will be positively correlated to pitch discrimination accuracy/sensitivity.

3) Determine the validity and reliability of the experimental stimuli (and online task environment) for assessment of perceptual-cognitive skills in youth and adult baseball players. I hope to show that the experimental task has high intra-person reliability and demonstrates validity in that it can distinguish across skill groups.

Chapter 2: Thesis

2.1 Participants

Male baseball players (n=21) were recruited via the University of British Columbia Baseball Centre, their affiliate organizations, and baseball organizations throughout Canada. We intended to collect data from a minimum n=15 per group, based on the research of Chen and colleagues (2017). However, our original aim was to collect data from as many eligible participants as possible, particularly at the skilled level, across a range of ages.

In addition to UBC baseball, 36 baseball organisations across Canada were contacted using information available on Baseball B.C. (www.baseball.bc.ca) and Field Level (www.fieldlevel.com). Organisations were filtered by age and level of play, and the head coach or technical director was contacted by email to provide information about the study. Players aged 12 to 22 years were targeted in recruitment and there was no discrimination based on player position (such as catcher, pitcher etc). We recruited experienced baseball players (more than 3 seasons playing experience) with the primary focus on determining age group differences (i.e. players of all skill levels were to be recruited). Information was collected via Qualtrics regarding the extent of an athlete's experience (number of years played, level of competition). Experienced participants included national, provincial, and regional representative players and those playing in the Premier Baseball League (PBL) or in the BC Minor Baseball. We also recruited participants who had played little to no baseball, or watched baseball regularly, but were not currently participating in organized baseball (i.e., minimal experience group). This latter group served as an adult novice control group (n=15). Novices were recruited via the UBC Psychology Paid Participants' Studies list and classroom advertisements. To ensure that participants were genuine, only enquiries from legitimate email addresses using the correct email format indicated in the advertisement were considered. All potential participants were required to confirm their age and experience with baseball via email before they were sent the link to the experience questionnaire. Whether the information provided in the email response matched the Qualtrics' questionnaire was checked to ensure legitimacy, and any participants who did not meet this requirement were excluded before being sent the Gorilla link to the main test stimuli.

The research was conducted according to the guidelines of (and with approval from) the University of British Columbia Behavioural Research Ethics Board (H19-02705; online study; H21-03828, in-person testing).

2.2 Stimuli

Three baseball pitchers from the college level were filmed to create the visual stimuli for the experiment. One pitcher's video was used for familiarization only, while the other two were used for the experiment. Prior consent was obtained from the pitchers. Rapsodo's Portable Pitching Monitor (Rapsodo, 2019), a unit which provides extensive information about ball release and flight during pitches, was used to confirm the pitch type, final location of the pitch relative to the strike zone, and ball velocity (Aucoin, 2019; Rapsodo, 2019).

Pitchers were asked to throw a series of pitches towards the strike zone of a hitter who would be approximately 175-180 cm tall. The supervising coach instructed the pitcher regarding pitch type and confirmed, in conjunction with the catcher and Rapsodo output, the type of pitch and final location relative to the strike zone (i.e. ball or strike). Each pitcher was asked to throw a maximum of 24 pitches. As illustrated in Figure 1, a Sony digital video camera was set up 354 cm back from the centre of the front of the home plate, and 64 cm to the left, with the lens at a height of 155cm. The frame rate was ~33ms. Camera and pitcher set up was based on the methods of Chen and colleagues (2017), as this simulates the perspective of an average height adult batter and ensures depth-cues are not lost by footage being taken from immediately in front of the pitcher.

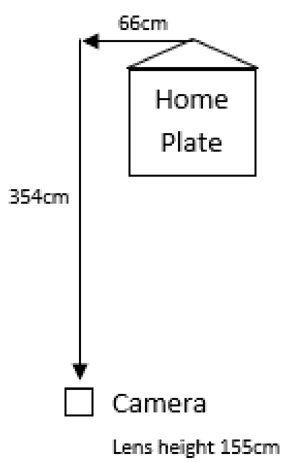
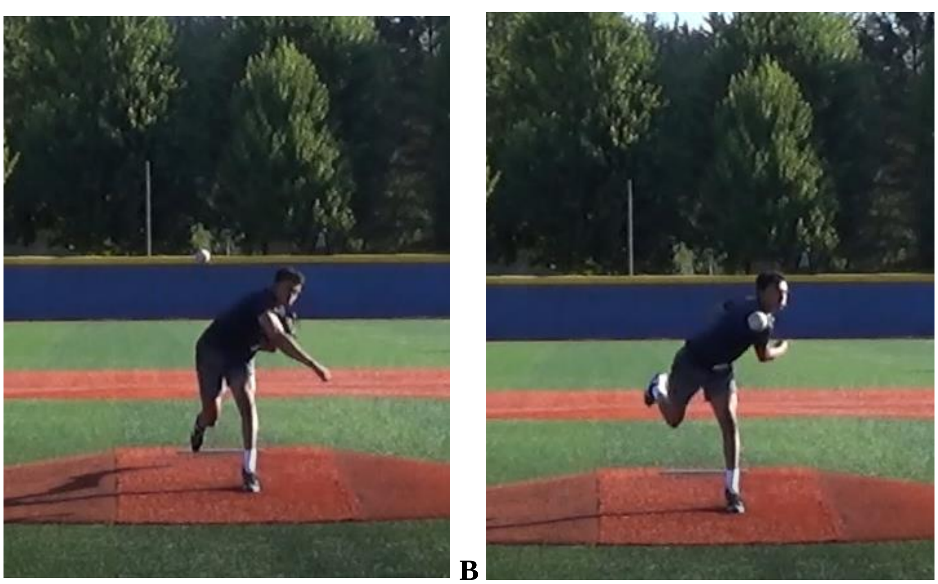


Figure 1. Set up for stimuli creation (note the camera would be behind the batter, if one was present).

Using Windows Video Editor, individual video clips were edited such that the outcome of the throw was occluded, with edits at various points in the unfolding action (see Figure 2). The first occlusion point (OP1) was made at 133 ms after the ball left the pitcher's hand. A duration of 133 ms more was applied to the edited video to create the next occlusion point (OP2) at 266 ms after ball release, and another 100 ms later at 366 ms for the final occlusion point (OP3). These time periods were based on the work of Chen and colleagues (2017). Brenton and colleagues (2016) showed that elite youth cricket batsmen were unable to predict bowl type above guessing until ball flight was shown supporting our decision to not include an “at-ball-release” occlusion point with these younger groups (see also Weissenteiner et al., 2008). The inclusion of relatively long clips helps to ensure that these younger athletes are able to respond to some of the videos with confidence.

We only asked participants to make predictions about pitch type. Morris-Binelli and colleagues (2018) showed that pitch location prediction accuracy was not above guessing for professional baseball batters until 200 ms after ball release. Further, combined pitch and location accuracy (asking for both responses at the same time) was not above guessing until the no occlusion condition (Morris-Binelli et al. 2018). We therefore did not expect that the younger participants (12-15 yr) or otherwise lesser experienced participants would be able to provide a confident response to pitch location before OP2 (266 ms after ball release).

Six videos were used from each of the two pitchers, representing 3 different pitch types (fastball, ~90mph; curve ball and change-up), such that there were 2 videos/pitcher for each pitch type (see Table 1). For the curveballs, we used only pitches which slightly missed the strike zone (termed balls). When reviewing the curveball pitches, the “strike” videos were not of high enough quality to be used.



A **B**

C 

Figure 2. Occlusion points: A) Ball release +133 ms (OP1). B) Ball release + 266 ms (OP2). C) Ball release + 366 ms (OP3). Pitcher A is shown.

Table 1: A breakdown of the number of clips as a function of pitcher, type of pitch, location of pitch and Occlusion Point (OP). These video clips were repeated during the task, so each clip was shown twice.

Pitcher 1						**Pitcher 2**						**Total**
Fastball		**Change-up**		**Curveball**		**Fastball**		**Change-up**		**Curveball**		
Strike 89mph (t=1)	Ball 89mph (t=1)	Strike 75mph (t=1)	Ball 78mph (t=1)	Ball 74mph (t=1)	Ball 77mph (t=1)	Strike 84mph (t=1)	Ball 86mph (t=1)	Strike 70mph (t=1)	Ball 72mph (t=1)	Ball 73mph (t=1)	Ball 73mph (t=1)	**12**
OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	OP 1-3 (t=3)	**36**

2.3 Procedure

2.3.1 Demographics and sports' experience online survey

Participants and/or parents provided assent and consent to participate by emailing the research team and completing an online survey with consent procedures included, via Qualtrics (UBC's recommended secure survey tool; https://ubc.ca1.qualtrics.com/jfe/form/SV_3TNC4feCgSqLAgK). In addition to details of the study, the survey was designed to elicit information about experience with baseball and other sports. Participants answered questions about their baseball playing experience including the positions played, years played and/or competed and level of competition. Month and year of birth, as well as current playing organization were collected to create age and skill groups. A minimum of three years of practice history data was collected (when available), to gain information about player history and determine reliability of collected information in subsequent years. The total time to complete the survey was ~15 minutes.

2.3.2 Baseball -specific online task

After completing the survey, participants were sent a link to the online task to complete via a laptop or desktop computer (with a code to enter, such that the experimental data contained no identifiable information without the code key). The online experiment was conducted using the Gorilla software platform for both design and administration (https://gorilla.sc/support). Participants responded via a key press to a series of occluded clips of either baseball pitcher. Each participant saw videos of the two pitchers and three pitch types, and they were asked to make decisions about pitch type. They were asked to sit as they made decisions in response to the videos and place the centre of the screen at eye level. Familiarization practice trials (t = 8; representing early, mid, and late OPs, 3 different pitch types, and location) were included to help individuals understand the requirements and to give some practice. Before the first phase of clips, feedback was provided as to the various types of pitches during this familiarization phase, as well as some instructions to explain the differences between the 3 pitch types. We included motivational feedback to encourage participants during the familiarization and experimental trials; including “good, keep going”, “well done, you’re doing great” etc., that were independent of accuracy.

During the experimental phase, participants were asked to respond to the given clip in the same manner as with the familiarization trials, inputting their predictions with a key-press corresponding to the pitch type (i.e., fastball, curveball, change up). Each participant saw every clip (e.g. Pitcher A, fastball, 266ms) twice throughout the online task. Participants were also asked for their confidence in their predictions at the end of each block (six trials) (from 0-100%, where 0 = not at all confident and 100 = extremely confident). Breaks were scheduled into the program between pitchers and the participant was able to control the break time (we showed clips of the same pitcher in blocks but the order of clips and pitchers was randomized across participants). The total time for this task to be completed was approximately 15 minutes, depending on the amount of rest between blocks.

We collected data on response time and accuracy, using baseline measures from familiarization trials, to standardize response times across participants based on computer set-up and keyboard delays. Participants were encouraged to respond as quickly as possible at the end of each trial without sacrificing accuracy.

2.3.3 Dynamic visual acuity task

After completing the prediction task, participants were asked to complete a short (~10 min), dynamic visual acuity (DVA) task using the same Gorilla software platform for administration of online experiments. We used a custom programmed moving Landolt C task, which required participants to determine the location of the opening of the “C” which moved across the screen at high speed (approximating baseball speed), with gap sizes ranging from one pixel (hardest) to four pixels (easiest) (Palidis et al., 2017; see also Laby et al., 2019 & Uchida et al., 2013). Participants responded by pressing a key signalling the location of the opening; either left-up, left-down, right-up or right-down.

The intention was to have this complete experimental protocol run as a longitudinal study, although this longitudinal aim was not part of my proposed Masters’ thesis. As such, the test and materials have been designed in such a way that the information can be solicited from players on a yearly basis (for ~3 years).

2.4 Analysis

The analyses have been broken down based on questions related to individual variables, rather than running large omnibus ANOVAs with 3-5 factors. The motivation for doing this was based on the fact that there were not significant numbers per age group to allow testing of the planned primary group measure. The current presentation of results is mostly motivated by a desire to evaluate the methods, particularly the discriminability and reliability of the test stimuli, and provide baseline data and recommendations for future research based on online tests of pitch recognition.

The plan was initially to focus data collection on skilled baseball players of varying ages. Due to issues with recruitment of the younger players, we mostly had data from players over the age of 18 years. As such, we recruited a group of novice, adult participants with limited to no experience with baseball, in order to show validity related to discriminability of the test stimuli. In terms of presentation of results, we first present the data from the skilled players (primary analysis) and then run secondary analysis comparing our skilled baseball players to a novice control group. The pitch discrimination data are presented before the dynamic visual acuity data and data are presented related to measures of proportional accuracy, d prime and response time.

Statistical analysis based on the skilled participants mostly involved paired t-tests or one or two-way ANOVAs; comparing across pitcher, pitch-type and/or occlusion point (all repeated measures' variables). When comparisons involved more than two means, Tukey's HSD post-hoc comparisons were made. Cohen's d is presented as a measure of effect size when comparing two means and partial eta squared for ANOVAs. Perceptual sensitivity analysis was conducted by way of d prime, based on hits (correct-detections) and false-alarms (incorrect detections). Hits and false alarms were corrected for skewness using a loglinear normalization procedure (as discussed in Hautus, 1995) as some participants were successful on all trials where a signal was present (1), or did not incorrectly 'detect' fastballs on noise trials where no fastballs were shown (0). Comparisons between fastballs and change-ups were the primary focus of analysis, as these two pitches are the most similar kinematically until late in ball flight. Between participant analyses were conducted to compare across player skill in comparisons of the expert and novice athletes (in mixed design ANOVAs). Correlational analyses (Pearson r) were used to relate measures of prediction accuracy to measures of DVA. Intra-person reliability analysis was conducted through calculation of % agreement.

2.5 Results

2.5.1 Baseball pitch discrimination task – Skilled baseball players

2.5.1.1 Pitch type discrimination

Participants were more accurate when predicting Pitcher B’s throws than Pitcher A’s, although there was considerable variability between players as shown in Figure 3. Pitcher B threw at a consistently lower speed (average 76 mph) than Pitcher A (average 80 mph) and had more traditional pitching mechanics, which may have contributed to the increased accuracy in prediction. Differences across pitchers was confirmed by a paired samples t-test, $t(20) = 6.20$, p=.001, Cohen’s d = 0.56. Because of differences due to pitcher, further analyses have been broken down by pitcher.

The average accuracy of pitch type prediction was above guessing (33 %) in all conditions except Pitcher A’s changeup. Accuracy was highest for fastballs, which may be related to the bias hitters have towards fastballs (see later sensitivity analysis) (Gray 2010; see also Loffing et al., 2015 for similar phenomenon in volleyball). Fastballs travel the fastest of the pitches, therefore hitters will often assume and prepare for a fastball until they receive information to indicate otherwise (Gray, 2002). Differences across pitch type were analyzed separately for each pitcher in one-way repeated measures’ ANOVAs. For Pitcher A, there was a significant effect of pitch-type, $F(40) = 19.50$, $p <.001$, $\eta_p^2 = 0.49$. Tukey HSD post hoc comparisons showed that fastballs were responded to more accurately than changeups ($p<0.001$) and curveballs (p = .002), and curveballs were responded to more accurately than changeups ($p = .045$). For Pitcher B, there was also a significant pitch-type effect, $F(40) = 76.10$, $p <.001$, $\eta_p^2 = 0.80$. Fastballs and curveballs were responded to more accurately than changeups ($p<.001$), but there was no significant difference in response accuracy between fastballs and curveballs.

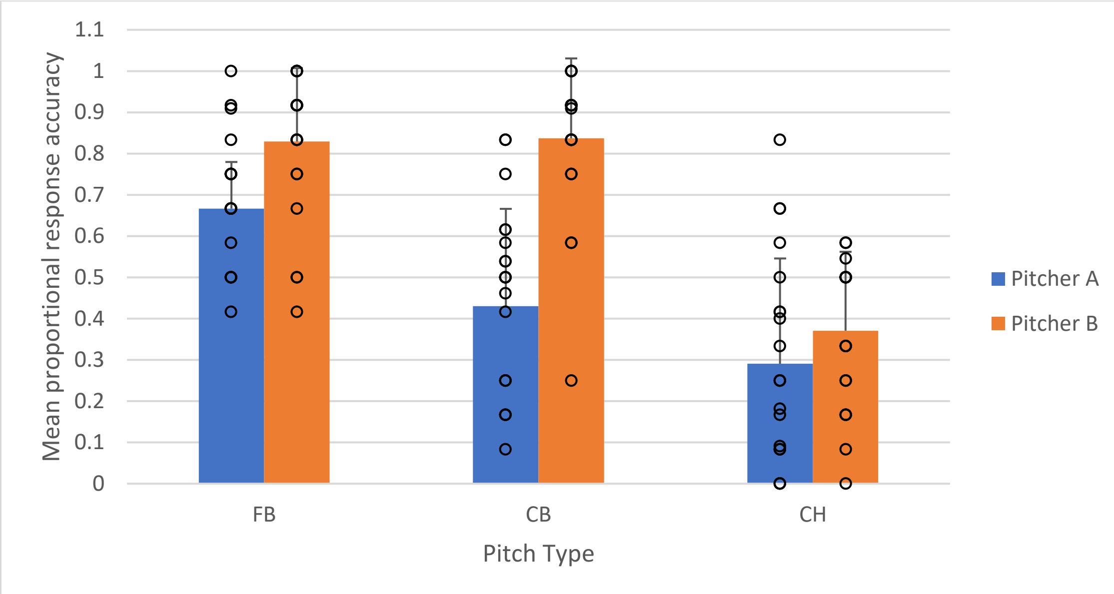


Figure 3. Mean proportional accuracy across pitch types: Fastball (FB), Curveball (CB), Changeup (CH), for pitchers A (blue bar) and B (orange bar). Error bars show SDs between participants. Data points show individual mean response accuracy.

2.5.1.2 Occlusion point

Accuracy was above guessing (33 %) at all occlusion points, but again dependent on pitcher as illustrated in Figure 4. Prediction accuracy did not improve as more information became available for Pitcher A, F (20) = 1.25, p = .30, η_p^2= .06. The prediction accuracy on Pitcher B's pitches was closer to what has been shown in previous anticipation studies, supporting the validity of these test stimuli (e.g., Chen et al. 2017; Müller et al., 2006). At 266 ms after ball release, there was enough information available for participants to identify pitch type at 70 % accuracy, but this did not improve further from OP2 (266 ms) to OP3 (366 ms). This Pitcher B effect was confirmed by a significant effect of occlusion point, F(20) = 10.20, p <.001, η_p^2 = 0.34. Tukey HSD post hoc comparisons showed significant differences in accuracy between 133ms and 266ms (p=.002) and 133ms and 366ms (p<.001), but not between 266ms and 366ms.

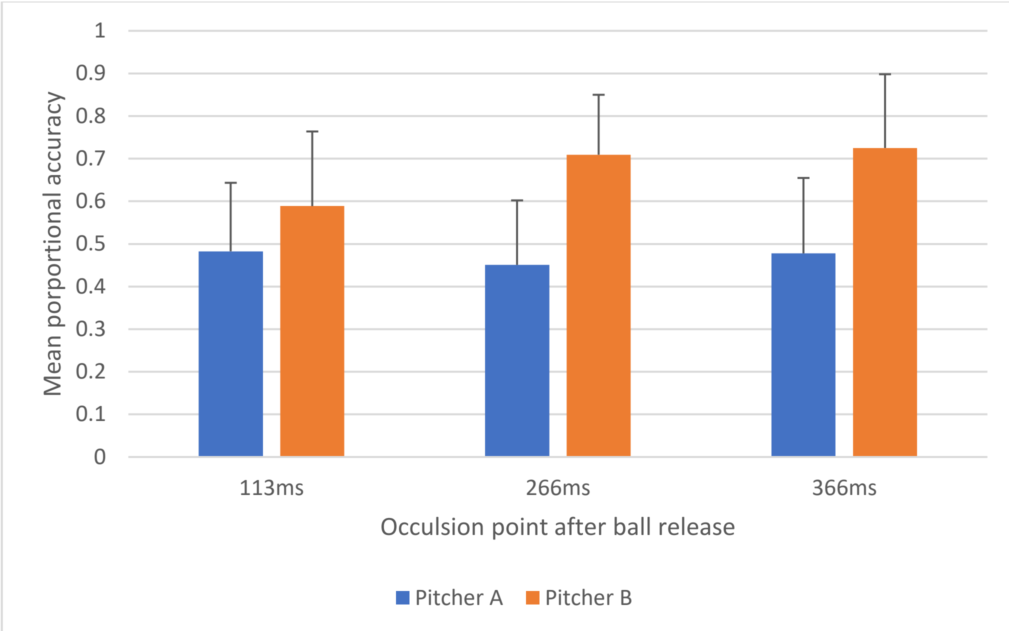
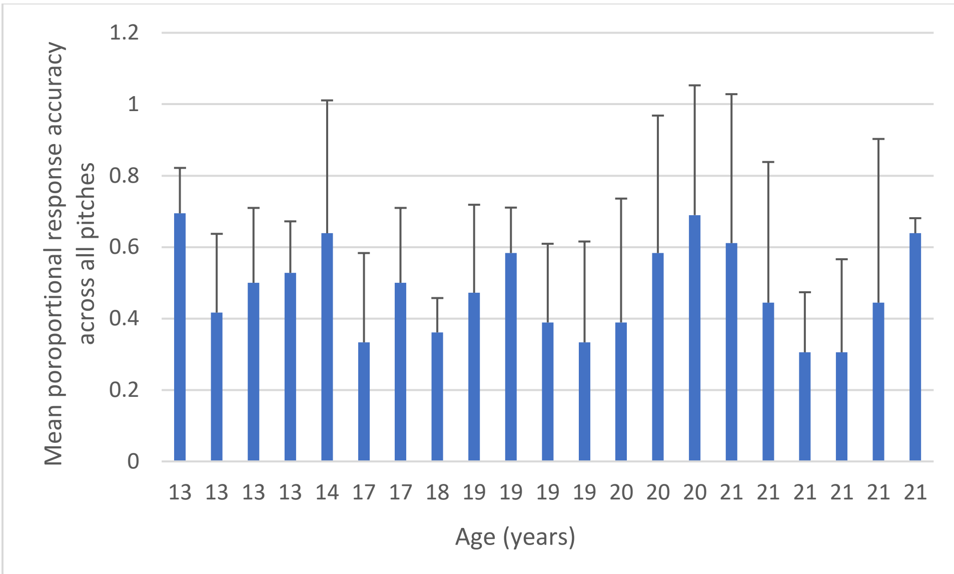


Figure 4. Mean proportional accuracy across all participants at each occlusion point for each pitcher. Error bars indicate SDs between participants.

2.5.1.3 Age effects

There was no evidence that age influenced prediction accuracy in this task, but there were only 5 athletes under the age of 15 years, with the majority of athletes being age 19-21 years. In Figures 5a-b, accuracy has been plotted for each individual participant labeled by age, for Pitcher A (a) and Pitcher B (b). One thing to notice was that there was considerable variation within an age group. We did not perform statistical analyses on these data.

a)



b)

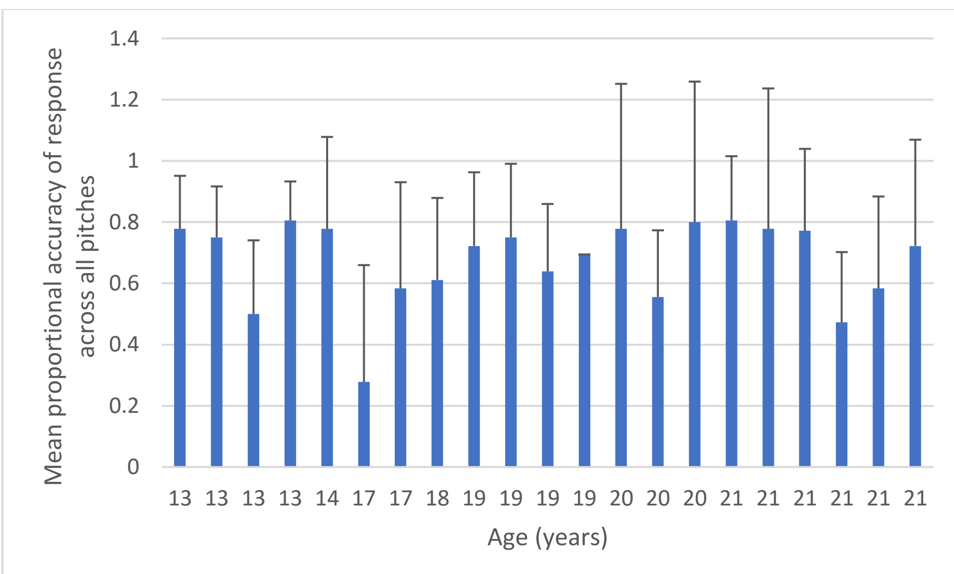


Figure 5. Mean proportional accuracy for each participant as a function of age, for Pitcher A (a) and Pitcher B (b). Error bars represent within participant SDs.

2.5.1.4 Player position

Players were assigned to a position category based on self-reporting during the Qualtrics' questionnaire, in which they identified where they spent most of their practice and playing time. 'Equal' indicates that the participant reported spending equal time practicing pitching and hitting. There were n =

6 Pitchers, n = 9 Hitters, n = 3 Catchers and n = 3 reported Equal time pitching and hitting. In general, there was little variation in prediction accuracy based on player position. Based on descriptive comparisons only, experience pitching resulted in slightly higher accuracy overall for Pitcher B, but this was not the case for Pitcher A (see Figure 6). Despite the fact that hitters have more visual experience seeing and responding to the thrown pitches, there were no trends for hitters to be more accurate than other player positions. Due to the uneven and low ns/group, we did not run statistical analyses on these data.

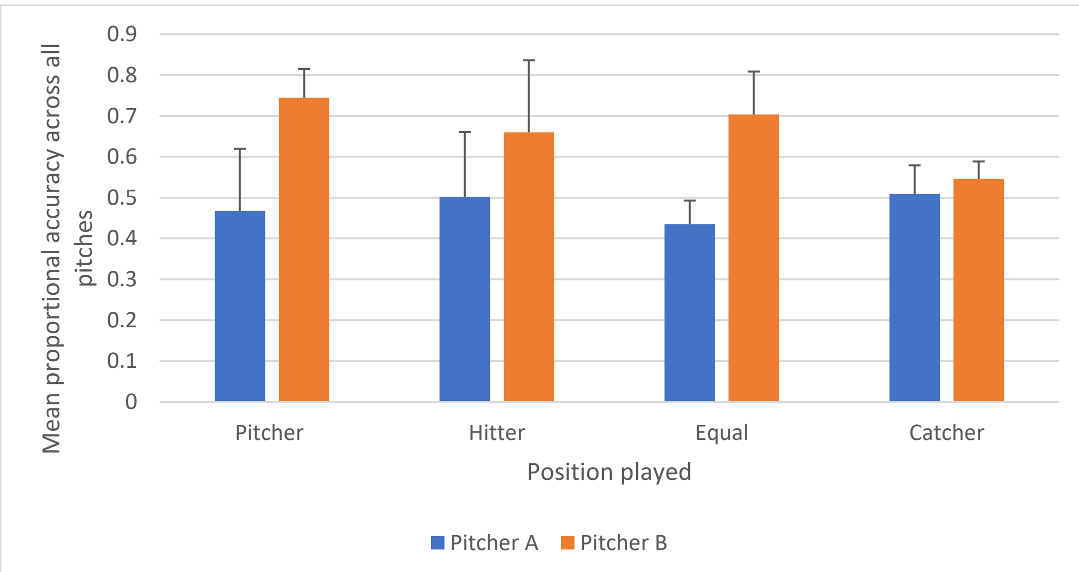


Figure 6. Mean proportional accuracy as a function of player position (summed across clips and pitch type). Error bars show SDs between participants.

2.5.2 Sensitivity analysis

2.5.2.1 Pitch-type discrimination

In Figure 7, frequency of hits and false-alarms are shown as a function of pitcher. Participants were more sensitive to the pitches thrown by Pitcher B (adj d' = 2.0) than to those thrown by Pitcher A (adj d' = 0.36), t(20) = 7.01, p<.001. The higher discrimination was a function of the high number of false alarms for Pitcher A in comparison to Pitcher B. Eighty percent of Pitcher B's fastballs were correctly identified (9.6 trials out of 12), and false alarms (incorrectly identifying a CH as a FB) occurred on 18% of trials (2.6 trials out of 12).

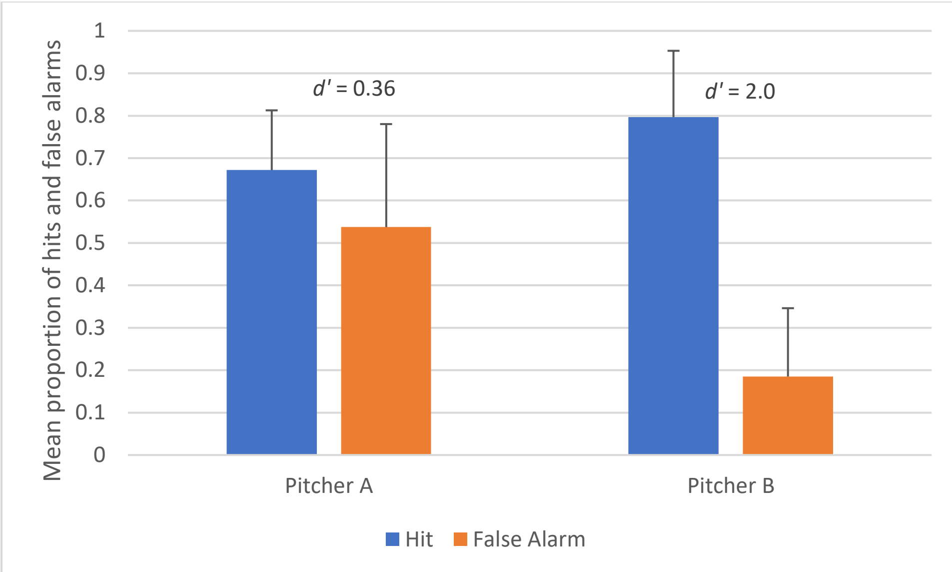


Figure 7. Mean proportion of hits (correctly identified a fastball) and false alarms (incorrectly identified a change-up as a fastball) as a function of pitcher. Adjusted *d* prime values shown. Error bars show SDs between participants.

2.5.2.2 Occlusion point

When sensitivity was broken down in relation to occlusion point for each pitcher, sensitivity to Pitcher A's fastball pitches was highest at the earliest occlusion point and decreased thereafter (Figure 8a). On these figures we have also provided average adjusted d prime values at each occlusion point. A one-way repeated measures' ANOVA comparing the individual log-transformed d prime values yielded a main effect of occlusion point for Pitcher A, $F(2, 40) = 6.06$, $p = .005$, $\eta_p^2 = .23$. Post hoc comparisons based on Tukey HSD confirmed the descriptive comparisons above, i.e. OP1 was significantly different from OP2 and OP3, which were not different from each other (*ps* < .05). In contrast, sensitivity to Pitcher B's pitches showed the predicted improvement from the earliest to the second occlusion and third occlusion points (Figure 8b). Despite these small differences in sensitivity, there was no significant effect of occlusion point, $F(2, 40) = 2.49$, $p = .96$, $\eta_p^2 = .11$.

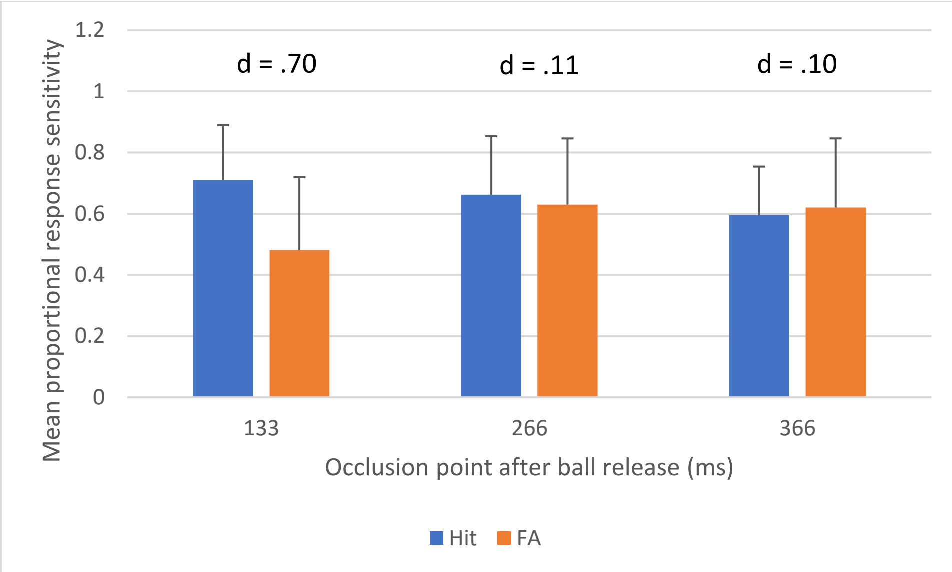


Figure 8a. Mean number of hits and false alarms and adjusted d prime values corresponding to fastball versus change-up detection at each occlusion point for Pitcher A. Error bars show between participant SDs.

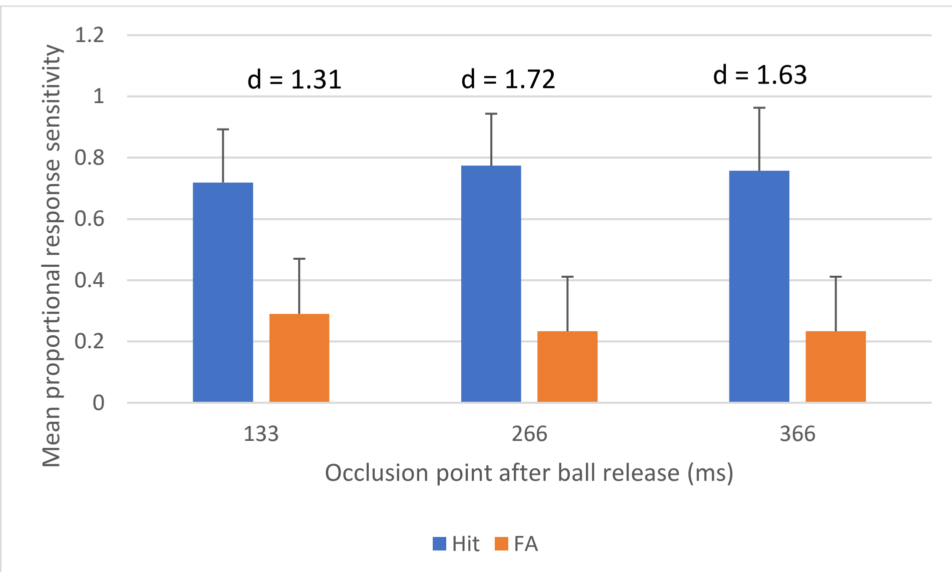
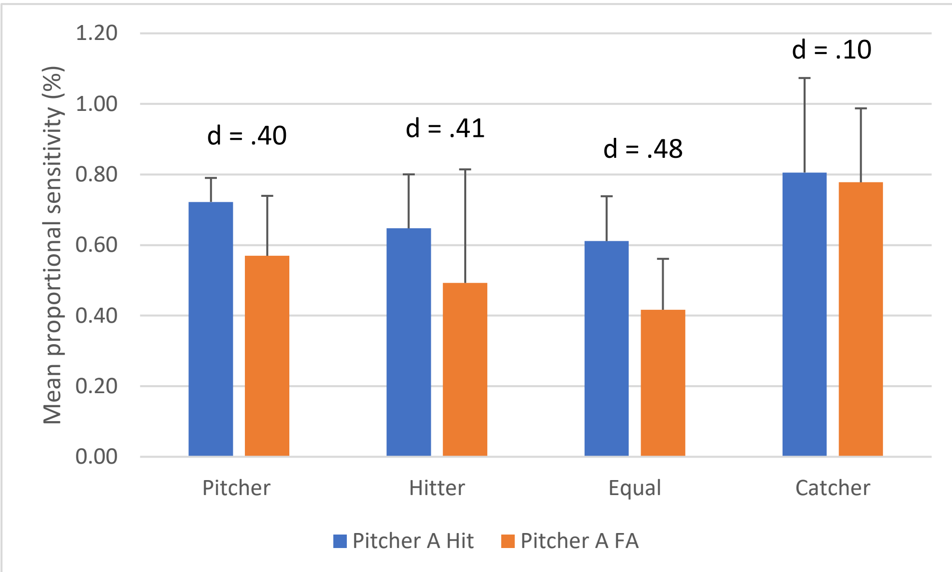


Figure 8b. Mean number of hits and false alarms and adjusted d prime values corresponding to fastball versus change-up detection at each occlusion point for Pitcher B. Error bars show between participant SDs.

2.5.2.3 Player position

Comparisons across hits, false alarms and d prime values for Pitcher A and Pitcher B are shown in Figures 9a and b as a function of player position. For Pitcher A, catchers were the least sensitive, based on the lack of differences between hits and false alarms (and hence low d prime scores), whereas there was little difference between the other player position groups (adding generally to test validity). However, for Pitcher B, pitchers (n = 6) were more sensitive to the fastball pitches, especially in comparison to catchers (n = 3) and hitters (n = 9) (Figure 6b). Because of the low power associated with the individual groups we did not run statistical analyses on these data.

a)



b)

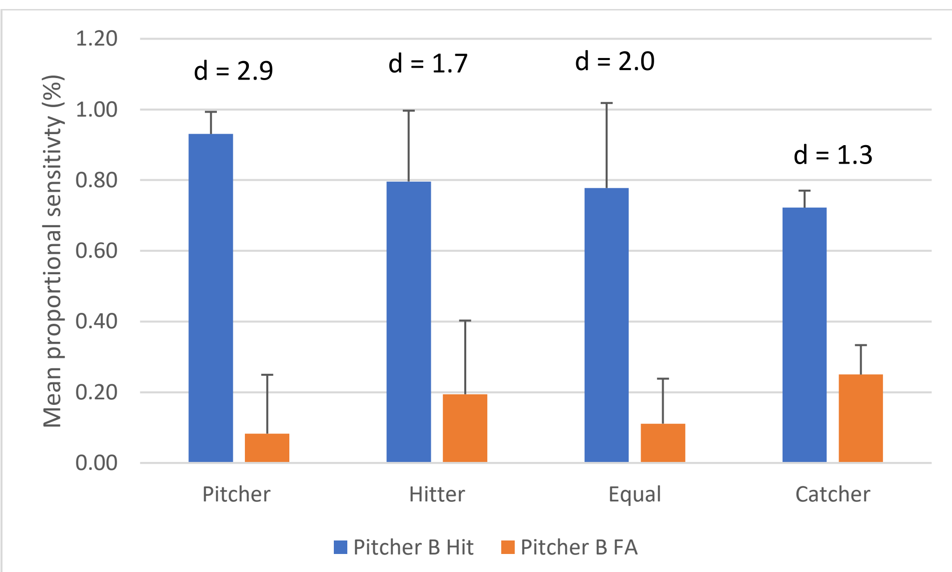
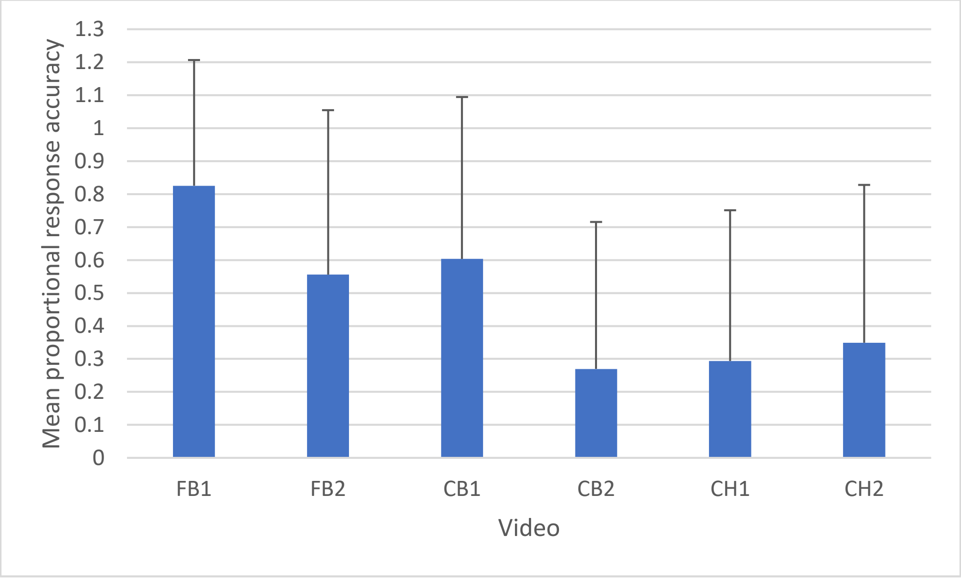


Figure 9a&b. Mean proportional sensitivity (adjusted d prime values) to fastballs (versus changeups) by position for Pitcher A (a) and Pitcher B (b). FA = False alarm. Error bars show SDs between participants.

2.5.3 Clip accuracy and reliability

The clips varied in their degree of difficulty, based on accuracy measures (see Figures 10a & b). For Pitcher A, the second FB and CB were responded to with lower accuracy than the first. Both the change-ups had low accuracy. For the more accurately predicted Pitcher B, the second change-up was clearly an anomaly and returned a prediction accuracy of 9% (equal to 0.54 trials correct out of 6/participant, or 11 trials out of 126 overall). Participants most often guessed that this change-up clip was of a curveball. Upon reviewing the clip, at ball release the pitch can be identified as a change-up because of the 'fat wrist' present in the pitcher's grip. Immediately after ball release, the ball appears to rise out of the pitcher's hand, which is what happens in the flight of a fastball (Bahill & Karnavas, 1993; Gray, 2010). However, towards the end of ball flight, the ball drops off, as a curveball would, hence the misidentification[1]. Inspection of accuracy as a function of occlusion point for this change-up video showed low accuracy at all occlusion points (i.e., OP1 = 0.07, OP2 = 0.12, OP3 = 0.07). This video should be removed from future testing and analysis.

a)



b)

[1] This should have highlighted a potential issue with this pitch based on responses from the catcher and coach present at stimuli collection, as both mistakenly identified this pitch as a change-up. It was also identified by the Rapsodo unit as a change-up.

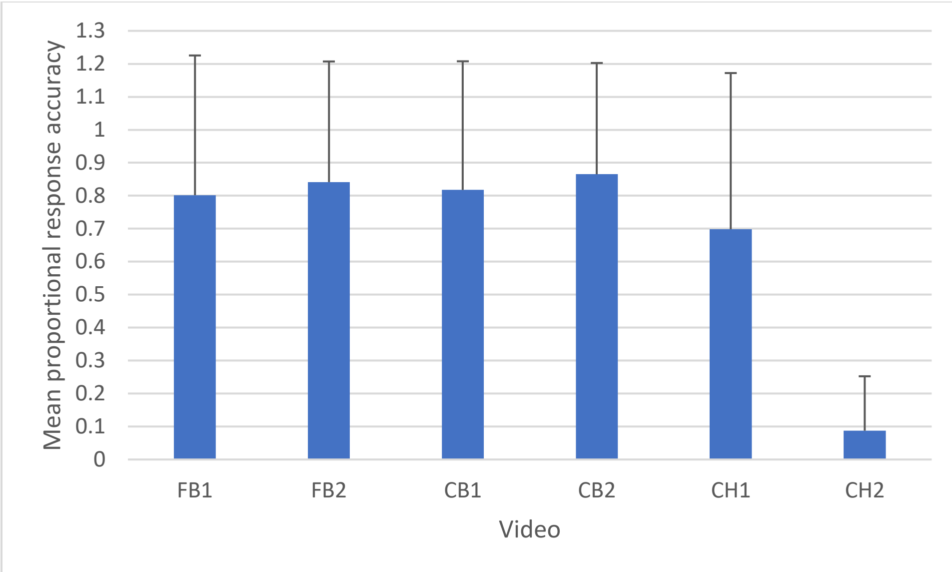
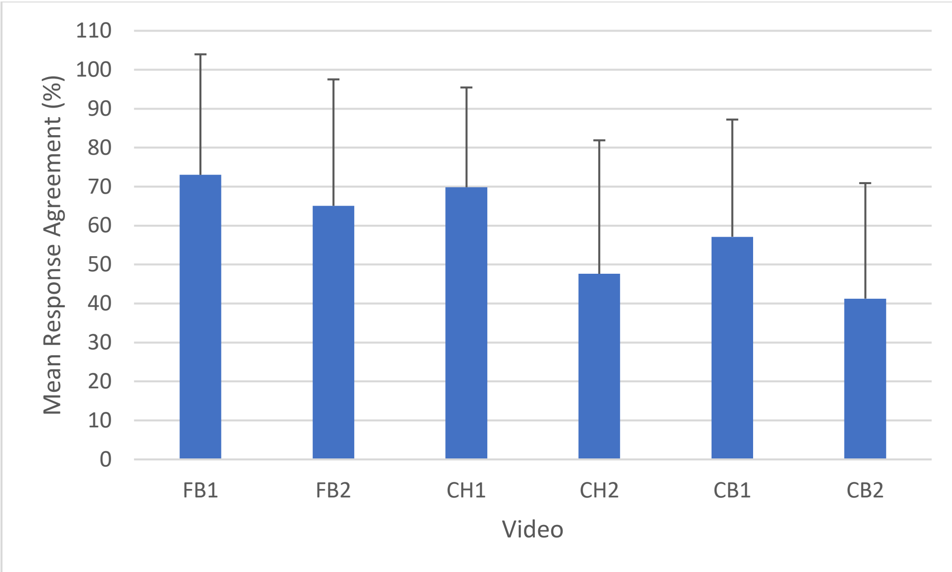


Figure 10a&b. Mean proportional response accuracy summed across skilled participants by clip for Pitcher A (a) and Pitcher B (b) (FB = Fastball, CB = Curveball, CH = Change-up). Error bars show SDs between participants.

Participants saw each clip twice through the experiment. Therefore, we were able to determine inter-trial (intra-person) reliability. To do this we calculated the percent agreement of responses to both exposures of each video clip. If participants gave the same response to both exposures, agreement was 100%, and if different responses were given, agreement was 0%. Agreement was calculated at each occlusion point and then averaged across all occlusion points for each participant. We have shown the average percent of agreement for Pitcher A and B in Figure 11a and b respectively, as a function of video clip (pitch-type). Looking at the means, reliability was lower for the videos of Pitcher A than for Pitcher B, particularly for change-up and curve-balls. Two of the videos of Pitcher A were responded to with the same response on both viewings less than 50% of the time on average.

We ran statistics on these data comparing percent agreement across the two pitchers and for each pitch type in a 2 X 3 fully repeated measures ANOVA. There was a main effect for pitcher, with higher agreement for Pitcher B than A, $F(1, 20) = 18.97$, $p < .001$, $\eta_p^2 = .49$. There was also a main effect for pitch type, $F(2, 40) = 3.36$, $p = .045$, $\eta_p^2 = .15$. Post hoc analysis of this effect yielded no significant differences between pitch types independent of pitcher. There was a Pitcher X Pitch-type interaction, $F(5, 100) = 9.98$, $p < .001$, $\eta_p^2 = 0.33$, with agreement for pitch type differing only for Pitcher A, not Pitcher B. That is, there was a significant difference between Pitcher A's curveball and fastball ($p = .05$), as well as between all of Pitcher B's pitches (*ps* <.006). Pitcher A's changeup was significantly different to Pitcher B's change up ($p = .002$) and curveball ($p = .02$).

a)



b)

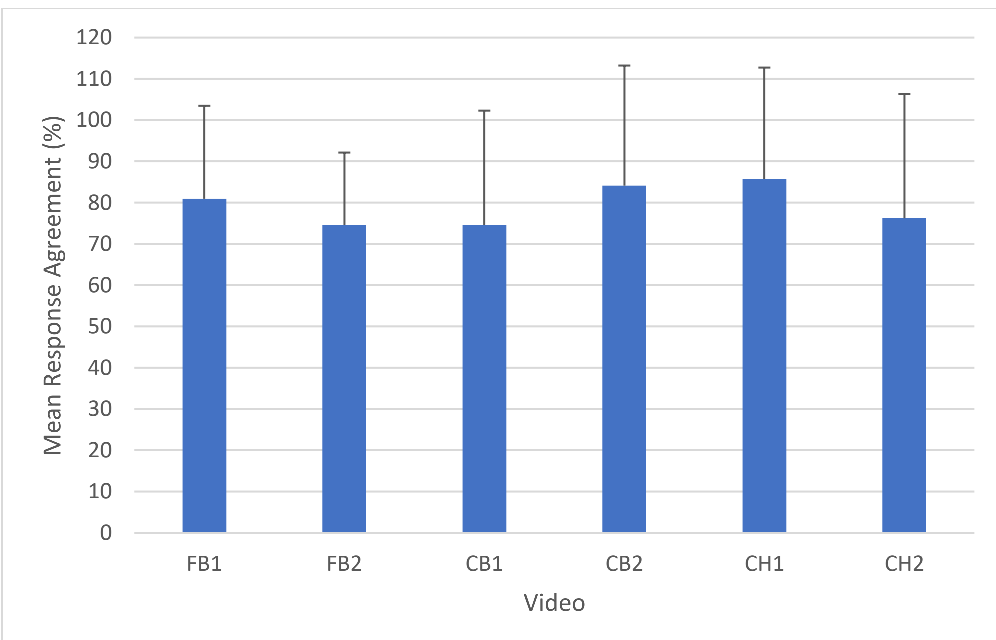


Figure 11a&b . Mean agreement in responses for each video clip as a function of pitch type (where FB = Fastball, CH = Changeup, CB = Curveball) for Pitcher A (a) and Pitcher B (b). Error bars show SDs between participants.

In Tables 2a and b, percent agreement on each video clip for Pitcher A and B as a function of occlusion point are shown. For both pitchers, percent accuracy showed small increases across occlusion points. Percent agreement of 81% (e.g. Pitcher A, FB 1, 133 ms) represents agreement by 17 out of 21 participants, while 71% agreement shows that only 14 participants agreed at the later (366 ms) occlusion point.

Table 2a. Video reliability (% agreement) by pitch type (FB = Fastball, CB = Curveball, CH = Changeup) and occlusion point (ms after ball release) for Pitcher A

Occlusion point: **Pitch clip:**	**133ms**	**266ms**	**366ms**	**Average:**
FB 1	81	67	71	**73**
FB 2	62	62	71	**65**
CB 1	62	38	71	**57**
CB 2	24	52	48	**41**
CH 1	62	76	71	**70**
CH 2	52	43	48	**48**
Average:	**57**	**56**	**63**	

Table 2b. Video reliability (% agreement) by pitch type (FB = Fastball, CB = Curveball, CH = Changeup) and occlusion point (ms after ball release) for Pitcher B

Occlusion point: **Pitch clip:**	**133ms**	**266ms**	**366ms**	**Average:**
FB 1	67	95	81	**81**
FB 2	76	81	70	**75**
CB 1	48	81	95	**74**
CB 2	71	90	90	**84**
CH 1	76	71	81	**76**
CH 2	81	86	90	**86**
Average	**70**	**84**	**84**	

2.5.4 Online vs. in person task

Six skilled online participants consented (through a separate ethics application) for us to use their data from an in-person version of this task, which was completed as a result of a subsequent UBC baseball testing phase. Only hitters participated in this stage of in-person testing. Briefly, the in-person version of the task was completed under the supervision of the research team. Participants stood in front of a TV screen, set up with their bat as though they were going to hit the incoming pitch, and responded verbally as quickly as possible with their pitch prediction. As with the online version, all clips were shown twice, for a total of 72 trials (*2 pitchers x 3 pitch types x 2 viewings of each pitch x 3 occlusion points x 2 trials each*). Somewhat surprisingly, the athletes were more accurate when completing the task online (mean proportional accuracy = 0.53 vs in-person = 0.46), although this difference was not significant, $t(5) = 1.29$, $p = .25$. When we broke down average accuracy for each pitcher, , differences across the two formats were most notable for Pitcher B, $t(5), = 2.33$, $p =.067$ (online mean proportional accuracy = 0.66 vs in-person = 0.56), in comparison to Pitcher A (online mean proportional accuracy = 0.40 vs in-person = 0.39). Within task, inter-trial reliability was also higher for the online task (mean percent agreement = 0.72, vs in-person = 0.54), and this was statistically significant, $t(5) = 2.89$, $p = .034$. Average inter-test reliability was 62% across the two tasks. Reliability across the two tasks was calculated by first calculating individual reliability for the online task within a person and then comparing this with the responses given in the in-person format (see Table 3). Individual percentages were then averaged across the group. As you can see in the table, if a person had 100% agreement for the online task, their responses on the in-person version would lead to a % agreement of 0, 50 or 100% if no response matched, one response matched, or both responses matched respectively. If the responses were different for the online version, but the individual gave these same responses again in-person, they got a score of 100% agreement. This reduced to 50% if both responses were the same for the in-person version (i.e., one different from the online) or 0% if neither response matched. We hope to get consent from other athletes to do further analyses on these in-person data in the future.

Table 3. Calculation of individual inter-test reliability across the online and in-person tasks.

Online	Same response on both trials			Different response on both trials		
In person	Both responses the same, both	Different responses, one matched online	Neither response matched online	Both responses the same, one	Different responses, both matched online	Neither response matched online

	matched online			matched online		
Agreement %	**100**	**50**	**0**	**50**	**100**	**0**

2.5.5 Response time

2.5.5.1 Average response time

Response time data were collected to help in assessment of performance and to determine potential speed for accuracy trade-offs. Individuals were encouraged to make fast responses, but they were told to prioritize accuracy over response time. Participants were given 3000 milliseconds to respond to a given clip before the response screen timed out. A total of 11 trials across all participants (2 skilled, 9 novice) were excluded due to time outs resulting in no response given. In Table 4 we have shown response times as a function of condition (pitcher, pitch type and occlusion point). The response times did not generally vary as a function of pitcher or pitch type. However, as the clip length increased, response time, on average, decreased.

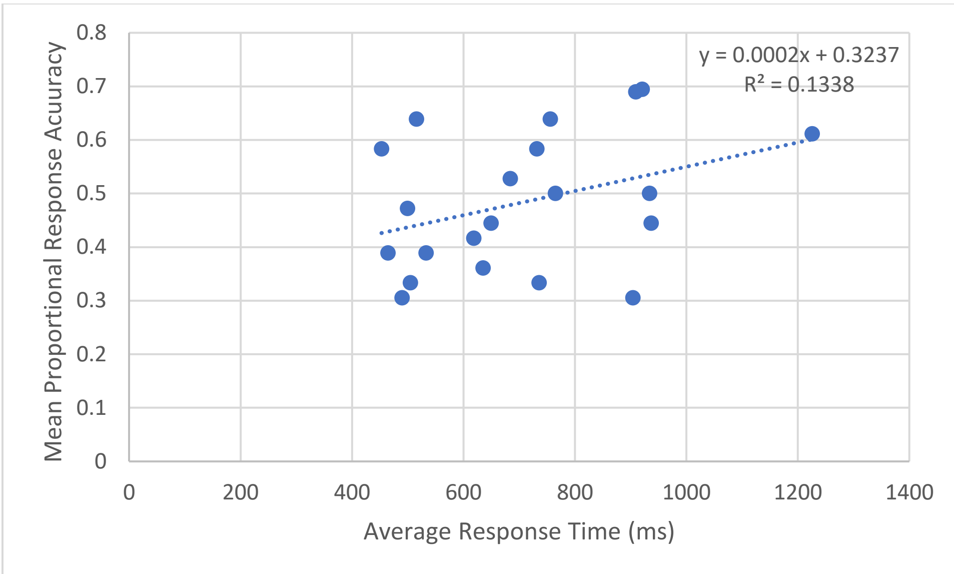
Table 4. Mean response times (ms) across all skilled participants for Pitcher A and B, for each pitch type and occlusion point.

	FB			**CH**			**CB**		
	133	**266**	**366**	**133**	**266**	**366**	**133**	**266**	**366**
Pitcher A	787.20	622.60	546.48	821.65	839.09	691.84	761.82	713.42	573.93
Pitcher B	703.93	515.97	451.71	601.74	574.53	411.78	744.92	466.81	389.47

2.5.5.2 Speed-accuracy relationships

There was a small-to-moderate size positive relationship between response time and mean accuracy for Pitcher A as shown in Figure 12a (r = .37, p =.09). Generally, slower responses were more accurate, but this was particularly notable for one participant whose mean response time was >1200ms. There was no relationship between response time and accuracy for Pitcher B (Figure 12b, r <.10).

a)



b)

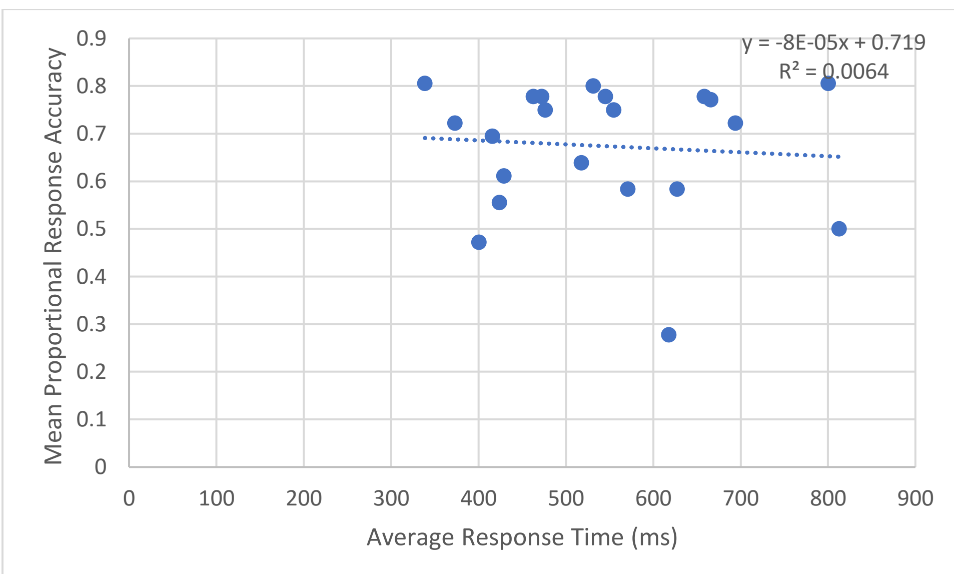
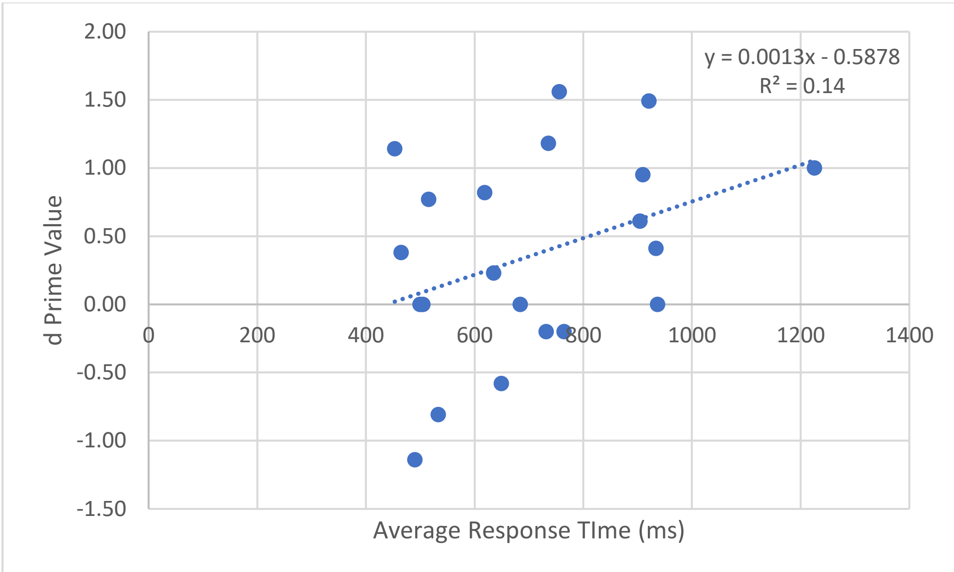


Figure 12a&b. Average accuracy for each participant compared to their average response time for Pitcher A (a) and Pitcher B (b). Trend line shows linear trend for response accuracy relative to response time.

Response time and response sensitivity (i.e., adjusted d prime) showed the same trend as noted for accuracy, with a positive but not significant trend between the two variables for Pitcher A (r =.37, p=.09, see Figure 13a), but no relation for Pitcher B (r<.10, see Figure 13b).

a)



b)

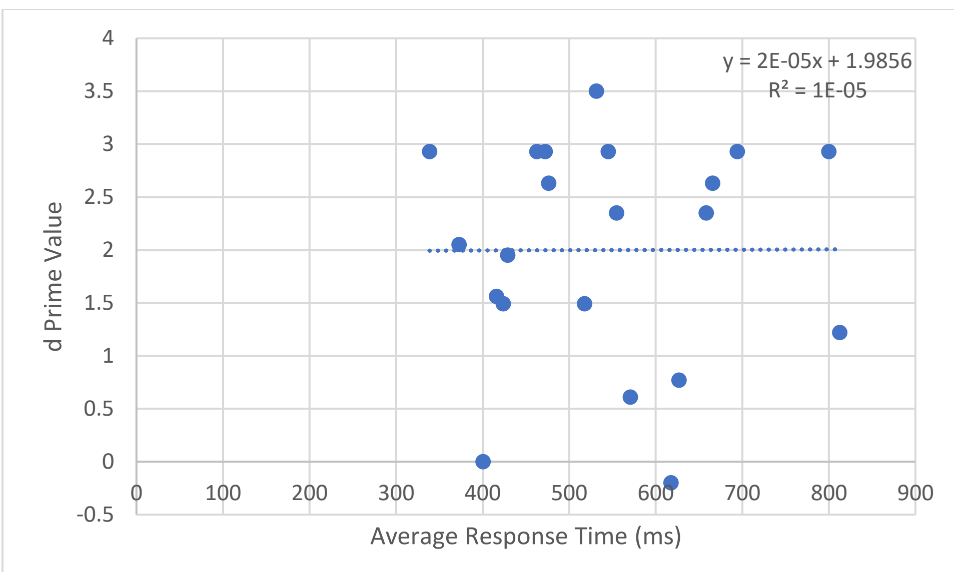


Figure 13a&b. Mean response sensitivity (d Prime) compared with mean response time for each participant for trials with Pitcher A (a) and Pitcher B (b).

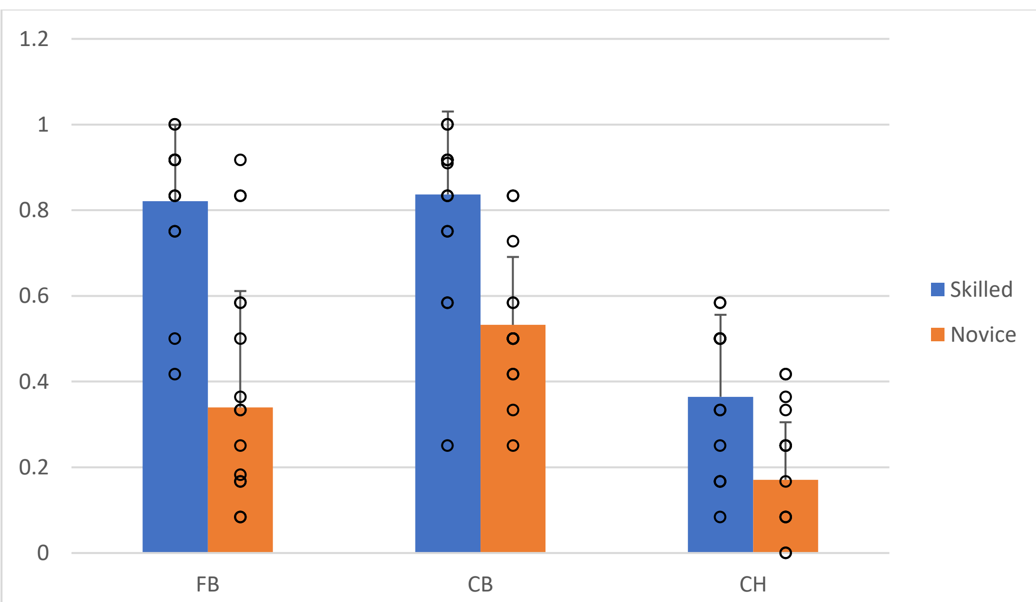
2.6 Novice group comparisons to the skilled baseball players

2.6.1 Prediction accuracy

2.6.1.1 Pitch type discrimination

As expected, the novice group had lower accuracy than the skilled group. A mixed design 2-way ANOVA (2 Skill X 3 Pitch-type), for Pitcher A, showed a main effect of skill group, $F(1, 34) = 19.50$, $p < .001$, $n_p^2 = .36$ and a main effect for pitch-type, $F(2, 68) = 13.87$, $p < .001$, $n_p^2 = .29$ (see Figure 14a). The Skill group X Pitch type interaction was not statistically significant, $F(2, 68) = 2.78$, $p = .06$, $n_p^2 = .70$. For Pitcher B, again there was a main effect of skill group, $F(1, 34) = 35.40$, $p < .001$, $n_p^2 = .52$, and pitch type, $F(2, 68) = 62.72$, $p < .001$, $n_p^2 = .66$, as well as a Skill group X Pitch type interaction, $F(2, 68) = 5.77$, $p = .005$, $n_p^2 = .15$. Tukey HSD post hoc analysis showed that the novice group performed significantly worse on all pitch types than the skilled group (*ps* < .05), but that the interaction was due to the fact that the for the skilled group, fastballs and curveballs were both responded to more accurately than change-ups, but for the novices, only the curveball was different to the change-up (see Figure 14b).

a)



b)

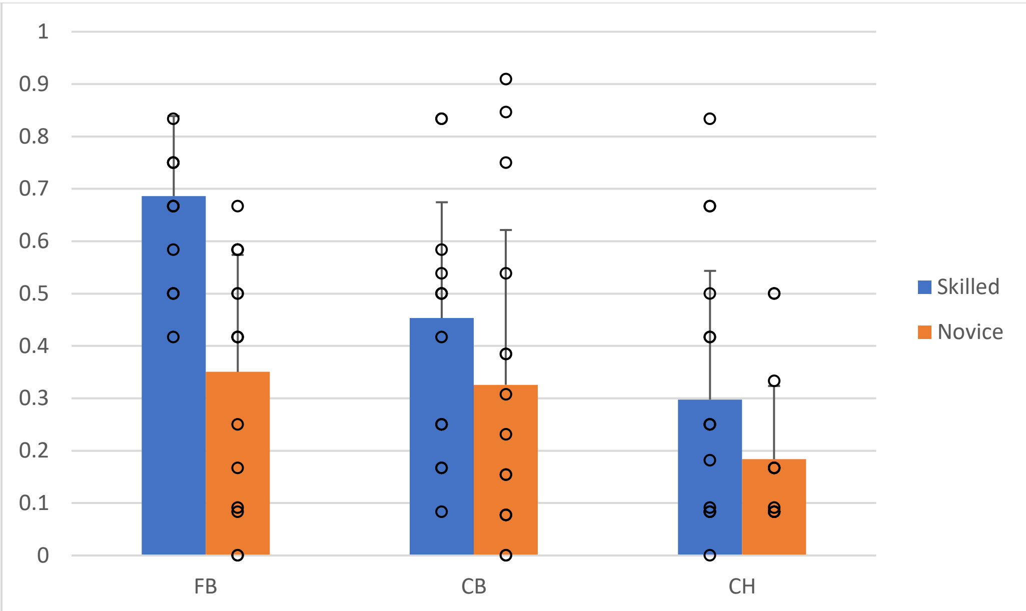
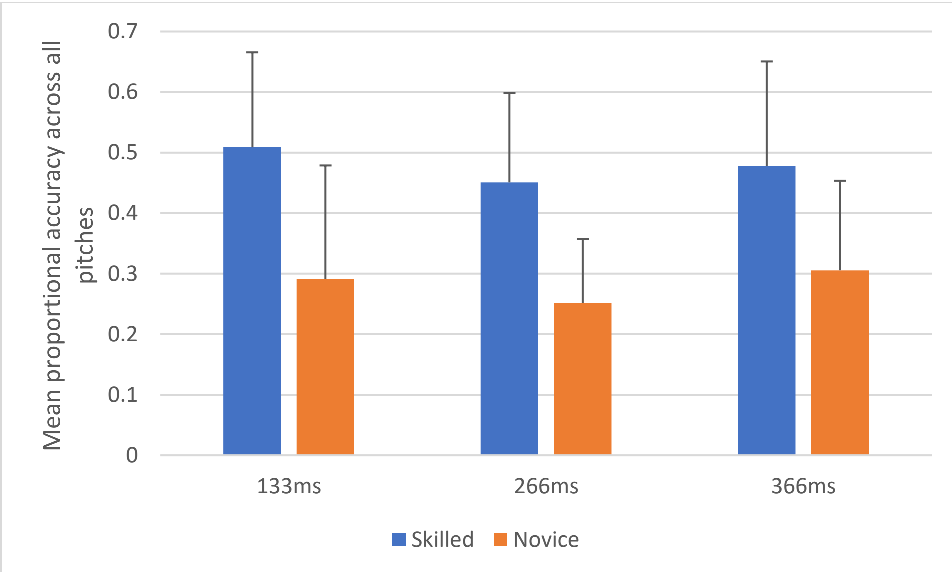


Figure 14a&b. Mean proportional accuracy for skilled versus novice participants for Pitcher A (a) and Pitcher B (b) (summed across clips). Error bars show SDs between participants. Data points show individual mean response accuracy. Note: There is overlap in results between individual participants.

2.6.1.2 Occlusion point:

Novices performed below guessing on average (33%) at all occlusion points for Pitcher A (Figure 15a), and at or just above chance for Pitcher B (Figure 15b). Separate analysis of Pitcher A, showed a significant skill group effect, $F(1, 34) = 20.30$, $p < .001$, $n_p^2 = .37$, but no significant effect of occlusion point ($F(2,8) = 1.74$, $p = .18$) or interaction between Skill group and occlusion point ($F<1$). For Pitcher B, there was again a significant effect of skill group, $F(1, 34) = 37.0$, $p < .001$, $n_p^2 = .52$, and occlusion point, $F(2, 68) = 10.92$, $p < .001$, $n_p^2 = 0.24$, but there was no interaction, $F(2,68) = 2.12$, $p = .13$.

a)



b)

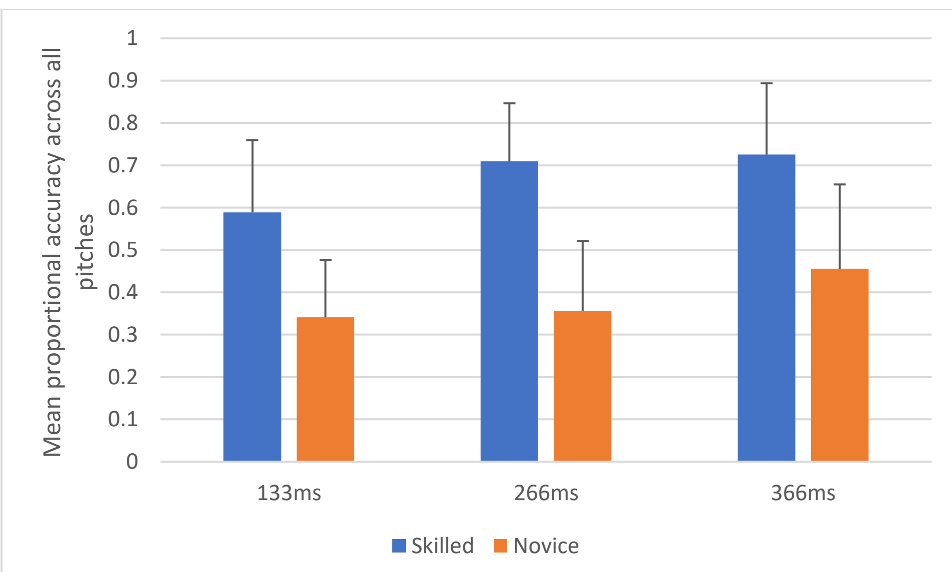


Figure 15a&b. Mean proportional accuracy at each occlusion point for the skilled and novice groups for Pitcher A (a) and Pitcher B (b), across all pitches. Error bars show between participant SDs.

2.6.2 Sensitivity

Analysis of log transformed d prime values showed that novices were significantly less sensitive to the presence of fastballs in fastball and change-up trials than experts for Pitcher A, t(34) = 2.07, *p* =

.046, Cohens d = .71 (mean adj d prime skilled = 0.21 novices = 0.12,) and Pitcher B, t(34) = 4.54, p < .001 (mean adj d prime novices = 0.45, skilled = 1.35). For Pitcher A, the novice group more often incorrectly identified change ups as fastballs than they correctly identified when a fastball was present (adj d' = -0.2). Novices were more sensitive to the difference in pitches thrown by Pitcher B, however this was a relatively small difference (d' = 0.45).

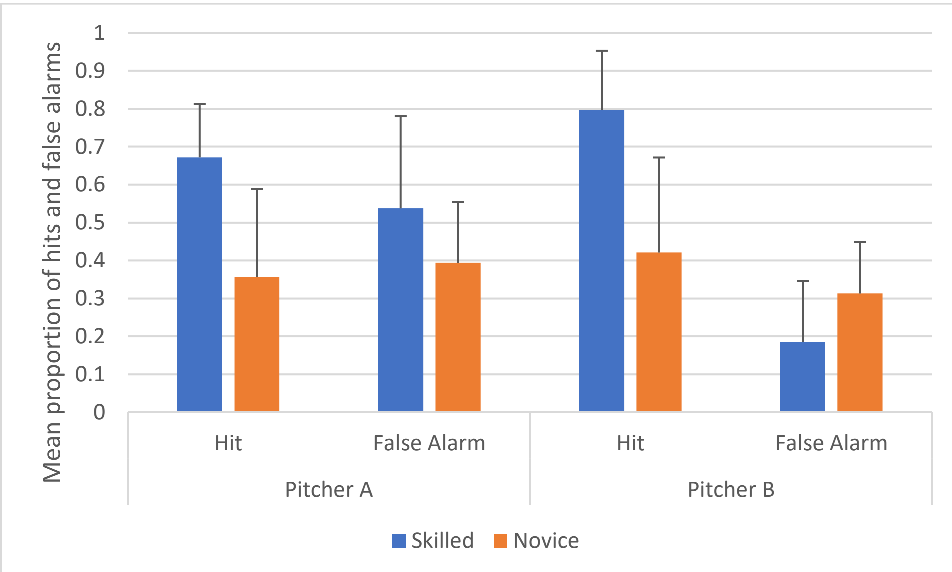


Figure 16. Mean proportion of hits and false alarms to fastballs versus changeups across skilled participants compared with novices for Pitcher A and B. Error bars show SDs between participants.

2.7 Dynamic visual acuity (DVA) task

2.7.1 Skilled baseball players

2.7.1.1 Mean proportional accuracy

In the DVA task (Landolt C), skilled participants' average accuracy decreased as the task difficulty increased (i.e. the gap in the Landolt C consisted of fewer pixels). Accuracy in terms of proportion of trials correct is displayed in Figure 17; with the highest difficulty noted as "1" pixel and easiest as "4" pixel gap size. At all levels of difficulty, participants performed above chance (25%).

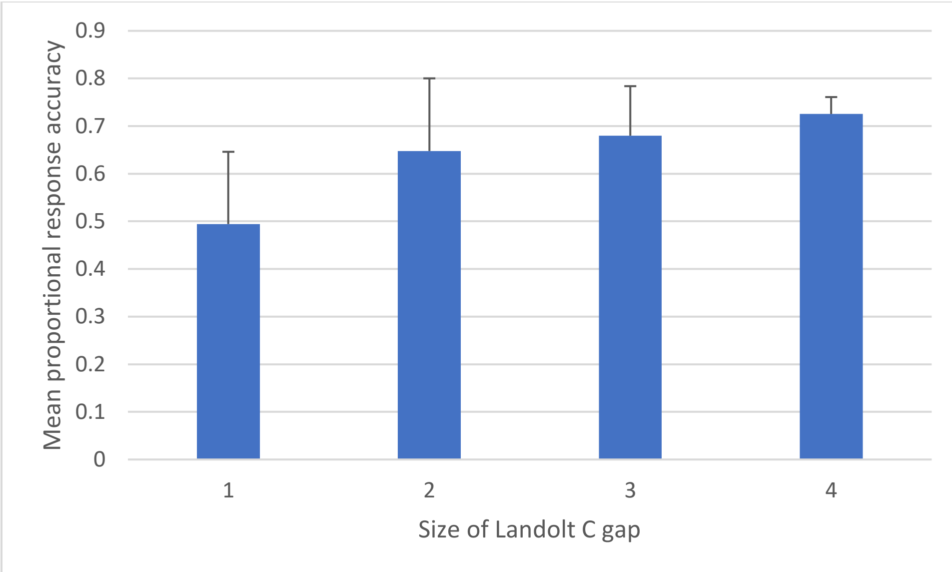


Figure 17. Mean proportion of trials correct at each condition of difficulty for skilled participants (where 1 = most difficult and 4 = easiest). Error bars show SDs between participants.

We hypothesized that there would be age group differences in the DVA task, however, as shown in Figure 18, there was no noticeable association between age and DVA. All individuals performed above guessing, but there was one individual in the 20 years age group who scored below 50% accuracy on this task.

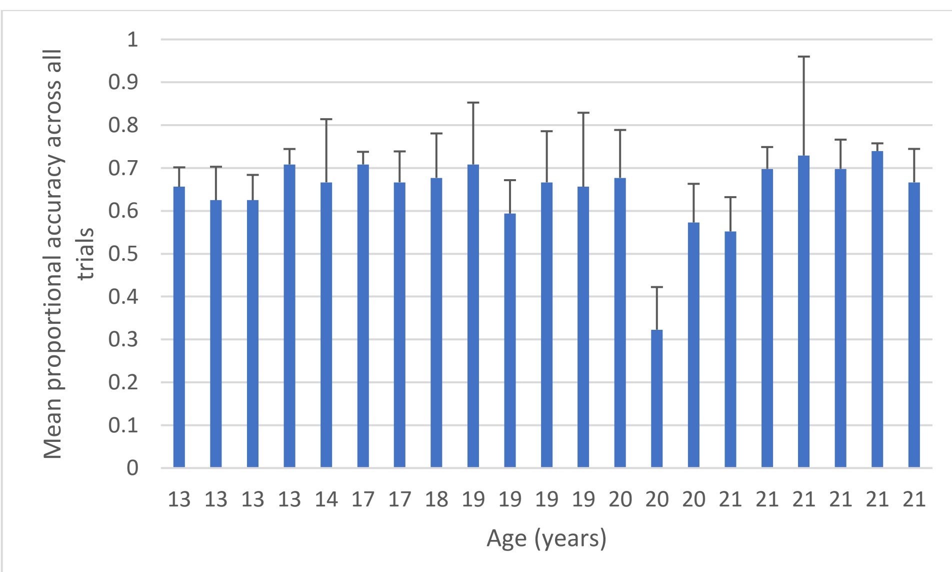


Figure 18. Mean accuracy on DVA trials by age for skilled participants. Error bars show SDs within participants.

2.7.1.2 Novice group comparisons to the skilled baseball players

We hypothesized that there would be skill group differences on the Landolt C task. These data were analysed in a 2 Skill group X 4 Gap size mixed design ANOVA. Although there was a trend for skilled participants to perform more accurately than the novice group as shown in Figure 19, this was not statistically significant, $F(1, 35) = 2.34$, $p = .14$, $n_p^2 = .06$. There was a significant effect of gap size, $F(3, 105) = 48.90$, $p < .001$, $n_p^2 = 0.57$. Tukey post hoc tests showed that this effect was due to the fact that the smallest gap (1 pixel) was significantly less accurate ($p < .001$) than all other gaps (2 – 4 pixels). There was no interaction between Skill group and Gap size, $F<1$.

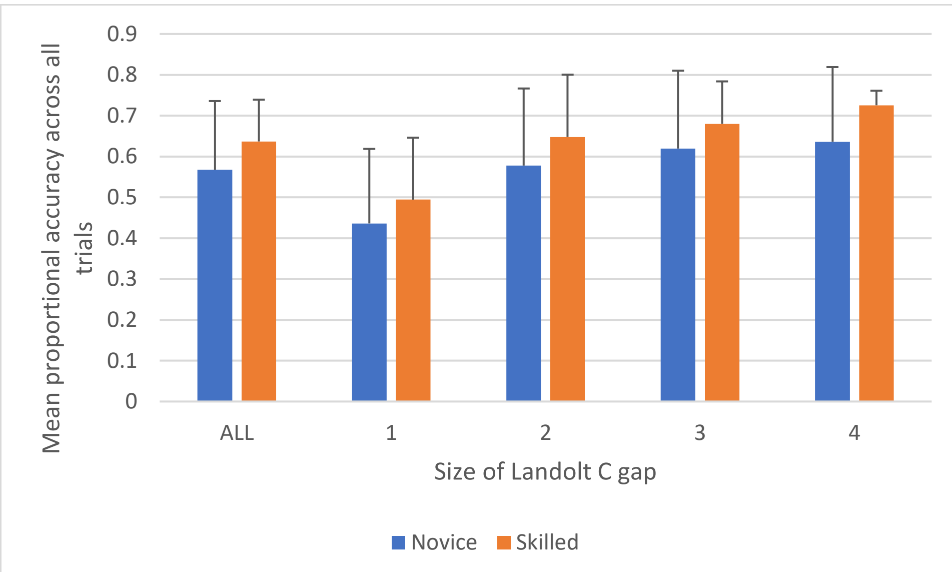
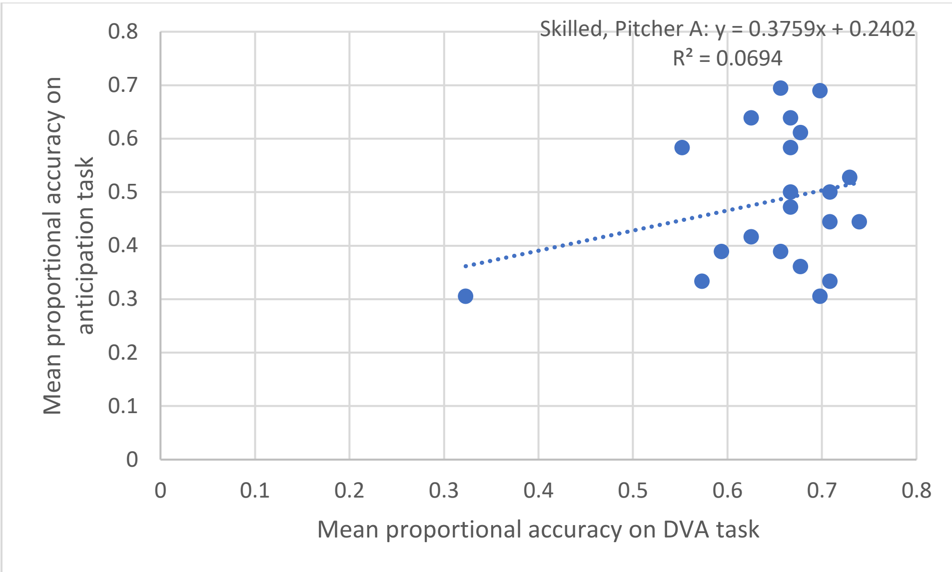


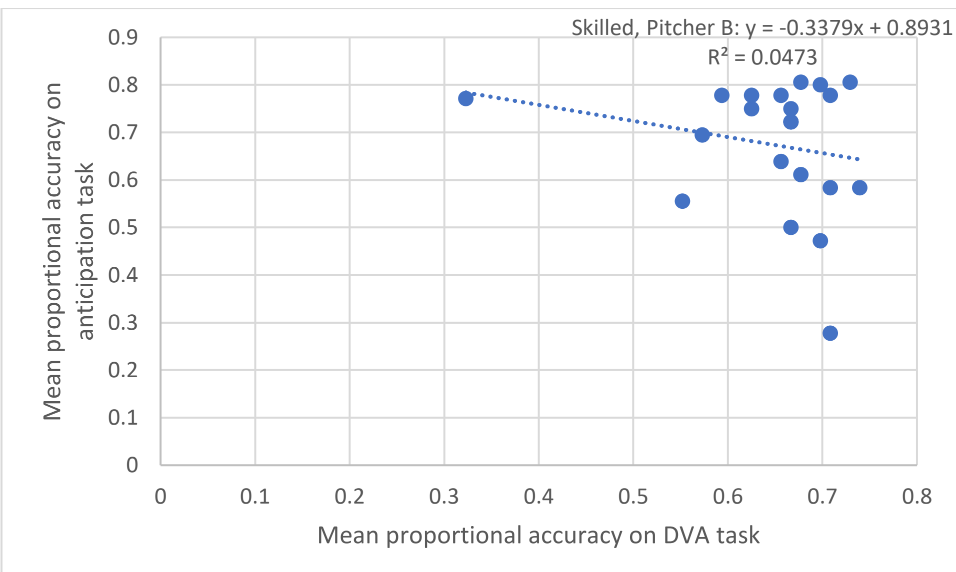
Figure 19. Mean proportional accuracy on DVA trials across skill groups. Error bars show between participant SDs.

For skilled participants, there was a non-significant trend for a positive relationship between average accuracy on the prediction task and DVA accuracy (r =.26, p=.25 for Pitcher A and r =.22, p=.34 for Pitcher B), as illustrated in Figures 20a and b. For novices, there was a statistically significant relationship between prediction task accuracy and DVA accuracy for Pitcher B (r =.59, p = .021), but not for Pitcher A (r =.41, p = .25), see Figure 20c and d.

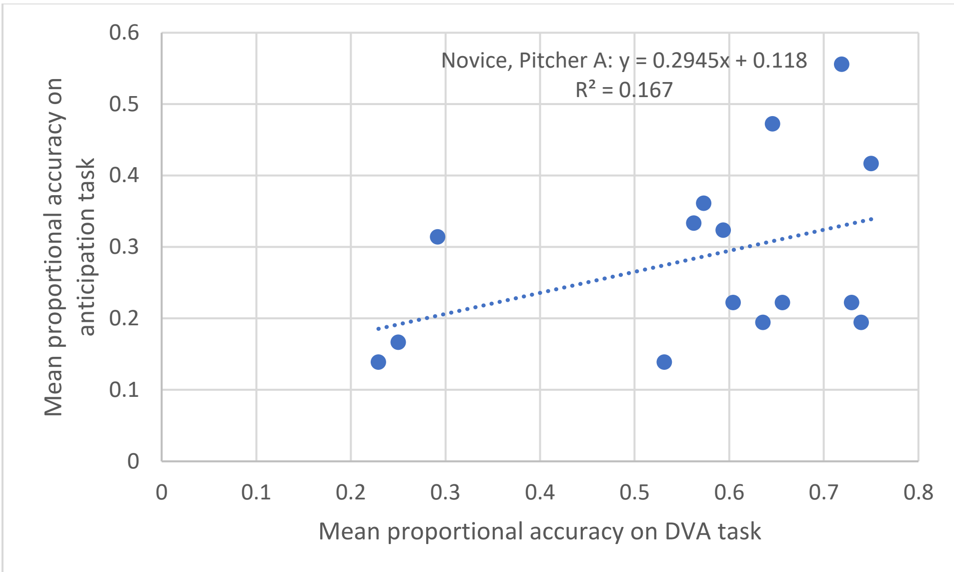
a)


Skilled, Pitcher A: y = 0.3759x + 0.2402
R² = 0.0694
Mean proportional accuracy on anticipation task
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Mean proportional accuracy on DVA task

b)


Skilled, Pitcher B: y = -0.3379x + 0.8931
R² = 0.0473
Mean proportional accuracy on anticipation task
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Mean proportional accuracy on DVA task

c)



d)

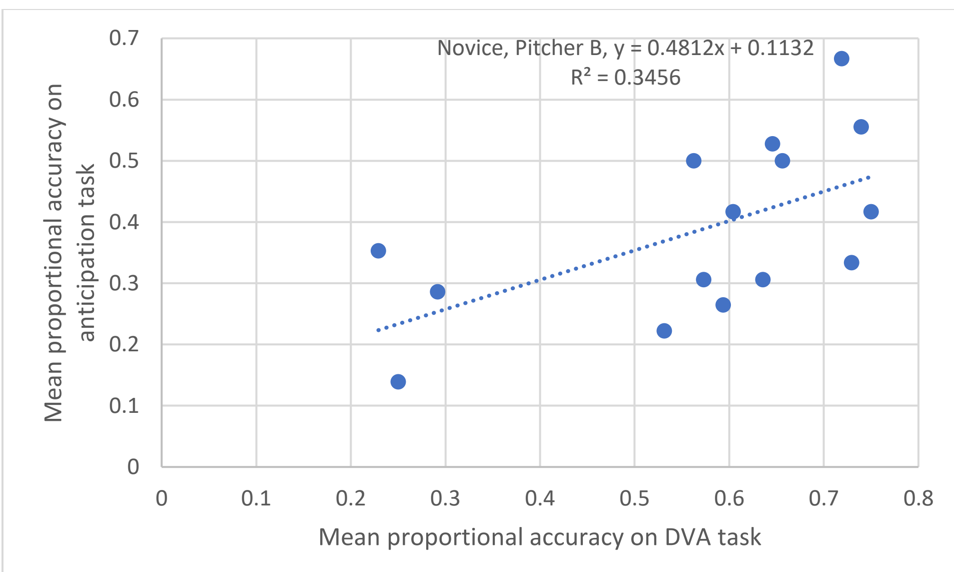


Figure 20a-d. DVA score compared with response accuracy for skilled (a & b) and novice (c & d) participants on trials with Pitcher A (a & c) and Pitcher B (b and d). Trend lines indicate linear relationships.

Chapter 3: Discussion and Conclusions

3.1 Summary statement of results

In general, we saw high accuracy, sensitivity, between-group discriminability (validity), and reliability for Pitcher B video clips, with accuracy being over 70%. The clips of Pitcher A generally returned less reliable and accurate results than those of Pitcher B (with the exception of the second change-up video). The differences across Pitcher may be a result of different pitch mechanics and speeds, with faster speeds for Pitcher A. Pitchers and hitters both performed well at this task, especially as compared to the catchers. We did not have sufficient representation from youth athletes to make any strong statements about age and pitch discrimination. From the data we did collect, there was no evidence of age-based individual differences in responding. We only saw a trend for skill-group differences in the Dynamic Visual Acuity (DVA) task. The novices showed more coupling between response accuracy on this task and their performance on the prediction task than the skilled participants.

3.2 Reliability

The mean agreement on clips of Pitcher A did not reach above 70%, whereas the clips from Pitcher B were consistently above 70%. Percentage agreement was higher on fastballs than other pitches, which was expected due to the assumed default response of expecting a fastball. For the clips of Pitcher A, the low reliability may indicate that the pitching mechanics and speed were too difficult for the participants in this sample watching a 2D video, and may suggest limitations in using 2D video to assess prediction accuracy for more unusual or fast paced pitches. The higher agreement for Pitcher B suggests that this pitcher's mechanics were easier to recognize and associate with a given pitch. The difference in pitch speed between pitchers likely aided here, as Pitcher B consistently threw 4-5mph slower than Pitcher A across all pitch types. We had a relatively high level of skill in our study sample, so we were surprised that the first pitcher was maybe too challenging for the athletes.

The percent agreement data showed that some of our current clips may not be appropriate to continue to use, as there was relatively low agreement across repeat viewings. This was most noticeable for Pitcher B's change up, at all occlusion points. Pitcher B had consistently higher agreement than Pitcher A, with agreement of 70% across all trials, except this one pitch. Pitcher A had lower agreement, with three pitches agreed on consistently less than 40% of the time (2.4 out of 6 trials agreed upon on average for each pitch, by each participant).

We found that the online task had higher agreement on clips than the in-person task. The agreement across all clips for both tasks was 61%, which means that there was agreement on 22 out of 36 clips, but mean agreement for the in-person task was only 55%. This may be due to the comparative difficulty of the in-person task, as participants were asked to set up with their bat as though they were attempting to hit the ball, and had to report their prediction to a member of the research team. We did not record response times for the in-person task, but we expect they would have been less than for the online task (mean response time for Pitcher A = 708 ms, Pitcher B = 542 ms) (see Ranganathan & Carlton, 2007).

3.3 Validity

In terms of the ability of the prediction test to discriminate across skill groups, there was good evidence of discriminability. The novice group performed worse than the skilled group. In line with previous studies, the novice responses did follow the expected model of increased accuracy as more ball flight information became available (e.g. Abernethy et al., 2001; Aglioti et al., 2008; Loffing et al., 2015). We did not, however, get any Skill group X Occlusion point interactions. There was no trend for larger differences between the groups for the early occlusion points (i.e., just after ball release), which may speak to differences in ability to use body kinematics from the pitchers to aid prediction accuracy. We acknowledge that the two groups in the present study were at near opposite ends of the expertise spectrum and so there is little nuance in our findings, at least with respect to within pitch discriminations. It will be important in future work to show that this task can discriminate between intermediate and highly skilled groups. The lack of age group differences suggests that there may be some issues in discriminating finer levels of experience/skill. Unfortunately, we were limited in the number of participants who comprised a younger age group with only n = 4 who were 13 years, and over half of study sample being over 19 years of age (n = 13).

Although we did not conduct a statistical analysis of player position, there was descriptive evidence of test validity shown in the positional breakdown of pitch discrimination. Catchers were less accurate overall than any other groups (pitchers, hitters, equal), and less sensitive to the fastball vs. changeup comparison. Catchers spend less time practicing hitting and pitching than other groups and watch ball flight from a different angle than hitters. As such, they would be comparatively lacking in both physical and visual expertise/experience. A descriptive comparison of pitchers (n=6), hitters (n=9), catchers (n=3), and those who reported being both pitchers and hitters (n=3), showed that pitchers performed with slightly higher accuracy than the other groups on the prediction task. These data do not align with those from Chen and colleagues (2017), who reported that batters were more accurate than pitchers on a location/strike-no-strike prediction decision. However, here we were specifically evaluating pitch discrimination not location, so it may be that the physical experiences of pitchers play a more significant role in such a pitch discrimination task. Due to time constraints in administering the task, we did not investigate final pitch location decisions. Moreover, Chen and colleagues only used clips of fastballs and used longer occlusion points (i.e., showed more of the clip), which may have also impacted their results. Hitters will usually hold a bias towards calling strikes because they are taught to plan to hit, so they are likely to be very accurate on strikes and often less accurate in identifying balls (Gray, 2010). Pitchers would be less likely to have this bias and so could perform comparatively worse on strikes. Again, acknowledging that we have only provided a descriptive analysis, our data fall more in line with the findings showing that athletes with motor experience in a specific skill perform better on similar action prediction tasks than those without such experience (e.g. Abernethy et al., 2001; Aglioti et al., 2008, Mulligan et al., 2016a; Urgesi et al., 2012).

It is difficult to know how ecologically valid our stimuli were and how closely the test environment would match decisions in the field. Based on analyses of Pitcher A, there may be cause to question its ecological validity. There is a pattern in the speed accuracy trade off results that showed participants, particularly skilled participants, improved their accuracy on Pitcher A trials if they slowed down and took longer to respond. As discussed in the introduction, baseball batters have less than 400 ms to respond to a pitch in a real game. As such, if these participants were to respond to Pitcher A in real life, they would have likely performed very poorly. In comparison, this pattern of increased accuracy with longer response times did not appear for Pitcher B. We allowed participants up to 3000 ms to respond, which was to ensure younger participants had a chance to respond to as many trials as possible, but this may reflect an ecological limitation. Moving forward, we may wish to reduce the response times available, in addition to using a different pitcher for stimuli. Given that the image is no longer available once the clip is finished, it is surprising that there should be advantages associated with responding later (at least for Pitcher A).

Data from an in-person version of this task adds to potential concerns about the ecological validity of online testing, where responses are not coupled to the stimuli (i.e. an actual physical response is required that is appropriate to the stimuli and coupled in time). In the in-person version of this task, we did ask participants to make a coupled response to the stimuli (i.e., to swing the bat) to help increase the validity. There is disagreement in the literature about whether action predictions need to be coupled for them to be valid measures of expertise. There is some evidence that expert accuracy is higher when the response is coupled (e.g. Farrow & Abernethy, 2003, Mann et al., 2010), yet there is also evidence that when they are uncoupled they are higher (e.g. Ranganathan & Carlton, 2007). Although we only had a small number of athletes complete both tasks, our results suggest that the uncoupled nature of the online task did not affect highly skilled prediction capability. Completing the prediction task in person where there was a higher level of action coupling did not lead to improved accuracy. Ranganathan and Carlton

(2007) proposed that the coupled task used in their baseball study (i.e. asking participants to swing their bat with the intention of intercepting a projected pitch) increased task difficulty as the response times were shorter (mean uncoupled response time = 814 ms, vs. coupled response time = 422 ms). We did not assess response times to the in-person stimuli, so we are unable to confirm whether this was a contributing factor to reduced accuracy. Given that response time and accuracy showed low correlations for Pitcher B for the online task suggests that it would not have been a significant contributing factor. However, testing of kinematics would be needed in future to better assess how the temporal demands associated with responding impact accuracy over and above any other attentional or pressure demands from in-person testing, associated with being watched and assessed.

3.4 Study Limitations

The results of the study were limited by the small sample size and unfortunately, despite plans, we were not able to statistically compare across age groups. The small sample size was a result of difficulties recruiting, which may be attributed to the online format of the study and fatigue associated with online tasks during the COVID-19 pandemic. There were also barriers associated with the necessary steps to get consent and assent as required by ethics, including third-party contact through coaches or teams, parental email correspondence and agreement through a Qualtrics' link, child assent and then finally, completion of a short Qualtrics' survey and the online tasks. Face to face testing may offer one solution to overcome these limitations, especially if the children are participating in week long camps. Maybe we need to gamify the environment more, but it seems that the task itself was not the impediment but the barriers to getting to the task. There was often a delay between the investigator sending the Qualtrics' task to participants, it being completed, and the Gorilla/main experimental prediction task being sent. In this time, it is possible that participants lost motivation to complete all the required tasks, despite us having a monetary incentive for completion. This drop out may not have occurred in person, as all tasks would be completed in one session in the presence of the investigator.

We acknowledge that the Qualtrics' task in addition to the Gorilla tasks may have made this experiment inaccessible to some participants. Of the 45 participants who received the Qualtrics' questionnaire, only 29 completed it, and of these 29, only 21 completed the Gorilla task. As discussed previously, this dropout may have been avoided if the task could have been run in person. Dropout may have also been avoided if the overall task was shortened, and the Qualtrics' questionnaire had been omitted. However, this would not have allowed us to collect practice history data, which was important to the original purpose of longitudinal tracking of children over time. Moreover, the initial phase of surveying participants on their practice history allowed us to maintain control over who participated to ensure internal validity.

3.5 Future directions

3.5.1 Stimuli

Based on the results above, new pitching footage needs to be obtained to replace some of the current videos. We have already recorded footage of another pitcher, which is currently being assessed for reliability in another study. We expect that there will be the opportunity to film pitchers later in the year, after the competitive season has been completed (April – May 2022). Because of COVID-19, we were limited at the time this experiment was launched to collect more video footage from pitchers. There is also a need to do a visual assessment of mechanics of our pitchers in addition to assessing pitch dynamics through Rapsodo, to assess whether the videos are representative of pitchers likely to be encountered by particular population and age groups. The data generated from this study reinforce the need to ensure footage used for stimuli is sensitive to skill group differences and reliably responded to over multiple viewings. These assessments and criteria will allow conclusions about the overall validity of 2D video footage (and video training apps) for games and assessments and whether such tests are a valid measure of skill in baseball. While we may be able to provide evidence to show that such tasks are able to discriminate across athlete skill levels, this does not justify whether or guide how individuals should use them as a mode of training. Fadde (2006) showed a relationship between in game statistics and use of a video-based training protocol with high level baseball players. However, this study did not show that the 'better' game statistics of the training group were a direct result of video training, nor that their statistics improved with regular and consistent training because a posttest-only design was used. As such, we still have a gap in the literature here that needs to be further investigated.

It will also be of interest to study left-handed pitchers in future stimuli creation, as these pitchers are less commonly faced but tend to present more of a challenge to right-handed batters. As such, this may present another level of difficulty and opportunity to assess perceptual cognitive skills.

The time allowed for responding was long and despite there being encouragement to respond to the stimuli as quickly as possible, we did notice some participants had timeouts on some trials. This was more prevalent in trials with novices, which suggests the trial may have been too difficult, or they may not have been paying attention. Time out trials were far less frequent among the skilled group, suggesting they were a) engaged, and b) understood the need to respond quickly. Knowing that the time allowed to respond was adequate, and that average response time was well within these maximum values, moving forward we may reduce the response time to increase difficulty and also provide a more realistic constraint.

3.5.2 DVA task and general perceptual skills

Based on past literature (e.g. Uchida et al., 2012, Klemish et al., 2018), there was reason to think that dynamic visual acuity (DVA) would differentiate across skill groups. Improved DVA may be a consequence of playing baseball, because the Landolt C task requires similar visual fixations and tracking to those required when watching a baseball (Laby et al., 2019; Uchida et al., 2012). We did indeed show some differences between novices and skilled participants on this task, but these differences were not statistically significant. Moreover, for the skilled athletes, there was only a weak relationship between performance on the DVA task and the experimental prediction task, but these relationships were stronger and statistically significant in novices. These differences may indicate that skilled participants were able to use early kinematic cues (e.g., grip, shoulder position) rather than ball flight, to better guide their predictions. Novices, however, were more reliant on ball flight (as evidenced in increased accuracy with longer clip lengths) which could be more demanding of dynamic visual acuity. Dynamic visual acuity is often referenced in regard to assessing athletes involved in high-speed interceptive sports (e.g. Burris et al., 2018; Klemish et al., 2018; Laby et al., 2019). Although it may be somewhat correlated with accuracy in a domain-specific prediction task, our data suggest that it is not a determining factor in action prediction accuracy for skilled players. This result is in contrast with the findings of both Laby and colleagues (2019) and Burris and colleagues (2018), who showed that dynamic visual acuity was positively associated with in-game statistics directly related to plate discipline (i.e. whether to swing or not). We did not compare DVA task results to in game statistics, showing an ecological limitation, nor was our skilled group of the elite level of the aforementioned studies' samples. In future, we may wish to use in-game statistics (e.g. swing rate, walk rate, on base percentage) to assist in assessing the validity of both the DVA task and the prediction task.

Klemish and colleagues (2018) have previously shown differences in youth baseball players and college and professional baseball players. Our DVA task did not appear to differentiate across the different age groups. Again, because of the low numbers of participants in each age group, particularly 12-14 years, it is difficult to make any strong conclusions about DVA and age. In future research, we hope to recruit more U15 participants, as well as track them longitudinally, to see if DVA changes with age and experience and how it tracks with playing experience as either hitters or pitchers.

3.6 Conclusions

In this study, we had initially sought to investigate how perceptual cognitive skills develop in youth athletes. However, due to difficulties in recruitment, this study evolved more into a test of the reliability and validity of perceptual-cognitive skills in skilled and novice baseball athletes through online assessment methods. There was evidence for test validity and reliability, mostly related to one of the pitchers and based on comparisons across skill groups, but not age. Skilled baseball players were significantly better at the prediction task than novices and showed more sensitivity to pitch type in their predictions, which is in agreement with past literature. Our skill groups were at extremes though, so we need to investigate an intermediate skill group using this task to assess whether a more nuanced discrimination is possible. The skilled group showed a trend for more accurate performance on the DVA task than the novice group, but this was not statistically significant. For the novice group, dynamic visual acuity was more related to performance on the pitch discrimination prediction task than for the skilled group. This may be due to several factors, including the reliance on late ball flight information by novices. For future research, there is a need to continue to refine our methods and stimuli based on further testing and analysis. We also need to consider ecological limitations with respect to the online task, as well as our recruitment methods to help accomplish aims of studying the development of perceptual skills in youth athletes longitudinally. We had hoped to make conclusions about the validity of perceptual assessments and video-based training apps, which will be better supported by data from youth athletes who are the target audience of such apps. Identifying the age that perceptual cognitive skills start to distinguish across skill groups will have implications for when such skills should be trained. Identifying methods for how to train these skills will be important for specific skill development and adding sport-relevant variety into training.

References

Abernethy, B. (1996). Training the visual-perceptual skills of athletes: Insights from the study of motor expertise. *The American Journal of Sports Medicine, 24*(6), 89-92.

Abernethy, B., Gill, D. P., Parks, S. L., & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception, 30*, 233-252.

Abernethy, B., Neal, R. J., & Koning, P. (1994). Visual-perceptual and cognitive differences between expert, intermediate, and novice snooker players. *Applied Cognitive Psychology, 8,* 185-211.

Abernethy, B. & Russell, D. G. (1987a). Expert-novice differences in an applied selective attention task. *Journal of Sport Psychology, 9,* (326-345).

Abernethy, B. & Russell, D. G. (1987b). The relationship between expertise and visual search strategy in a racquet sport. *Human Movement Science, 6*, 283-319.

Adler, D. (2019, Feb 24). *Votto's plate discipline in a league of its own: Reds star's approach at the plate has him set up for a big 2019*. MLB. https://www.mlb.com/news/joey-votto-still-has-mlb-s-best-plate-discipline

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11*(9), 1109-1116.

Aucoin, D. (2019). *Validating Rapsodo 2.0: Pitch design made easier*. Driveline Baseball. https://www.drivelinebaseball.com/2019/07/validating-rapsodo-2-0-pitch-design-made-easier/

Bahill, A. T. & Baldwin, D. G. (2004). The rising fastball and other perceptual illusions of batters. In: Hung, G. K. & Pallis, J. M. (eds). *Biomedical Engineering Principles in Sports. Bioengineering, Mechanics, and Materials: Principles and Applications in Sports.* Vol 1. Springer, Boston, MA.

Bahill, A. T. & Karnavas, W. J. (1993). The perceptual illusion of baseball's rising fastball and breaking curveball. *Journal of Experimental Psychology: Human Perception and Performance, 19*(1), 3-14.

Brenton, J., Müller, S. & Mansingh, A. (2016). Discrimination of visual anticipation in skilled cricket batsmen. *Journal of Applied Sports Psychology, 25*, 483-488.

Burris, K., Vittetoe, K., Ramger, B., Suresh, S., Tokdar, S. T., Reiter, J. P., & Appelbaum, L.G. (2018). Sensorimotor abilities predict on-field performance in professional baseball. *Scientific Reports, 8*(116).

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E. & Haggard, P. (2005). Action observation and acquired motor skills: An FMRI study with expert dancers. *Cerebral Cortex, 15*(8), 1243-1249.

Cañal-Bruland, R., Kreinbucher, C. & Oudjeans, R. R. D. (2012). Motor experience influences strike and ball judgements in baseball. *International Journal of Sport Psychology, 43,* 137-152.

Chen, Y., Lee, P., Lu, Y., Huang, S., & Yen, N. (2017). Contributions of perceptual and motor experience of an observed action to anticipating its result. *Journal of Experimental Psychology: Human Perception and Performance, 43*(2), 307-316.

Davids, K., Araujo, D., Vilar, L., Renshaw, I., & Pinder, R. (2013). An ecological dynamics approach to skill acquisition: Implications for development of talent in sport. *Talent Development & Excellence, 5,* 21-34.

Dicks, M., Button., C. & Davids, K. (2010). Examination of gaze behaviours under in situ and video simulation task constraints reveals differences in information pickup for perception and action. *Attention, Perception, & Psychophysics, 72*(3), 706-720.

Dokken, N. (2019, Jan 29). *Sabermetric series, part 3: Plate discipline*. FantraxHQ. https://www.fantraxhq.com/sabermetric-series-part-3-plate-discipline/

Fadde, P. (2006). Interactive video training of perceptual decision-making in the sport of baseball. *Technology, Instruction, Cognition and Learning, 4,* 237-256.

Farrow, D. & Abernethy, B. (2003). Do expertise and the degree of perception-action coupling affect natural anticipatory performance? *Perception,* 32, 1127-1139.

Farrow, D., Abernethy, B. & Jackson, R. C. (2005). Probing expert anticipation with the temporal occlusion paradigm: Experimental investigations of some methodological issues. *Motor Control, 9*(3), 330-349.

Farrow, D. & Reid, M. (2012). The contribution of situational probability information to anticipatory skill. *Journal of Science and Medicine in Sport, 15,* 368-373.

Fooken, J. & Spering, M. (2020). Eye movements as a readout of sensorimotor decision processes. *Journal of Neurophysiology, 123*(4), 1439-1447.

French, K. E. & McPherson, S. L. (1999). Adaptations in response selection processes used during sport competition with increasing age and expertise. *International Journal of Sports Psychology, 30*, 173-193.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualisations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review, 23,* 523-552.

Gray, R. (2010). Expert baseball batters have greater sensitivity in making swing decisions. *Research Quarterly for Exercise and Sport, 81*(3), 373-378.

Gray, R. (2017). Transfer of training from virtual to real baseball batting. *Frontiers in Psychology, 8*(2183).

Gray, R. (2019). Sports training technologies: Achieving and assessing transfer. In: Hodges, N. J. & Williams, A. M. (Eds), *Skill Acquisition in Sport: Research, Theory, and Practice, 3rd ed.* (pp. 203-219) London: Routledge.

Gray, R. & Canal-Bruland, R. (2018). Integrating visual trajectory and probabilistic information in baseball batting. *Psychology of Sport and Exercise, 36*, 123-131.

Hodges, N. J., Wyder-Hodge, P. A., Hetherington, S., Baker, J., Besler, Z., & Spering, M. (2021). Topical review: Perceptual-cognitive skills, methods, and skill-based comparisons in interceptive sports. *Optometry and Vision Science, 98*(7), 681-695.

Hommel, B., Musseler, J., Aschersleben, G. & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioural and Brain Sciences, 24,* 849-937.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G. Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with our own mirror neuron system. *PLoS Biology, 3*(3), 529-535.

Jones, C. M. & Miles, T. R. (1978). Use of advance cues in predicting the flight of a lawn tennis ball. *Journal of Human Movement Studies, 4*(4), 231-235.

Kato, T. & Fukuda, T. (2002). Visual search strategies of baseball batters: Eye movements during the preparatory phase of batting. *Perceptual and Motor Skills, 94,* 380-386.

Klemish, D., Ramger, B., Vittetoe, K., Reiter, J. P., Tokdar, S. T., & Appelbaum, L. G. (2018). Visual abilities distinguish pitchers from hitters in professional baseball. *Journal of Sports Sciences, 36*(2), 171-179.

Laby, D., Kirschen, D. G., Govindarajulu, U., & DeLand, P. (2019). The effect of visual function on the batting performance of professional baseball players. *Scientific Reports*, *9*(1).

Loffing, F., Stern, R. & Hagemann, N. (2015). Pattern-induced expectation bias in anticipation of action outcomes. *Acta Psychologica*, *161*, 45-53.

Mann, D. L., Abernethy, B. & Farrow, D. (2010). Action specificity increases anticipatory performance and the expert advantage in natural interceptive tasks. *Acta Psychologica*, *135*(1), 17-23.

Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport, 14*(17), 2135-2136.

Morris-Binelli, K. & Muller, S. (2017). Advancements to the understanding of expert visual anticipation skill in striking sports. *Canadian Journal of Behavioural Science, 49*(4), 262-268.

Morris-Binelli, K., Muller, S. & Fadde, P. (2018). Use of pitcher game footage to measure visual anticipation and its relationship to baseball batting statistics. *Journal of Motor Learning and Development*, *6*(2), 197-208.

Müller, S. & Abernethy, B. (2012). Expert anticipatory skill in striking sports: A review and a model. *Research Quarterly for Exercise and Sport*, *83*(2), 175-187.

Müller, S., Abernethy, B. & Farrow, D. (2006). How do world-class cricket batsmen anticipate a bowler's intention? *Quarterly Journal of Experimental Psychology, 59*(12), 2162-2186.

Müller, S., Fadde, P. J. & Harbaugh, A. G. (2017). Adaptability of expert visual anticipation in baseball batting. *Journal of Sports Sciences, 35*(17), 1582-1690.

Mulligan, D., Lohse, K. R., & Hodges, N. J. (2016a). An action-incongruent secondary task modulates prediction accuracy in experience performers: Evidence for motor simulation. *Psychological Research, 80,* 496-509.

Mulligan, D., Lohse, K. R., & Hodges, N. J. (2016b). Evidence for dual mechanisms of action prediction dependent on acquired visual-motor experiences. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1615-1626.

Muraskin, J., Sherwin, J. & Sajda, P. (2015). Knowing when not to swing: EEG evidence that enhanced perception-action coupling underlies baseball batter expertise. *Neuroimage, 123,* 1-10.

Nakamoto, H., Ikudome, S., Yotani, K., Maruyama, A. & Mori, S. (2013). Fast-ball sports experts depend on an inhibitory strategy to reprogram their movement timing. *Experimental Brain Research, 228*(2).

Palidis, D. J., Wyder-Hodge, P. A., Fooken, J., & Spering, M. (2017). Distinct eye movement patterns enhance dynamic visual acuity. *PLoS ONE*, *12*(2).

Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91,* 176-180.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*(2), 129-154.

Ranganathan, R. & Carlton, L. G. (2007). Perception-action coupling and anticipatory performance in baseball batting. *Journal of Motor Behavior,* 39(5), 369-380.

Rapsodo. (2019). *Pitching 2.0.* https://rapsodo.com/baseball/pitching/

Rizzolatti G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27,* 169-192.

Savelsbergh, G. J. P., Williams, A. M., van der Kamp, J. & Ward, P. (2002). Visual search, anticipation, and expertise in soccer goalkeepers. *Journal of Sports Sciences, 20,* 279-287.

Smeeton, N. J., Huttermann, S. & Williams, A. M. (2019). Postural cues, biological motion perception, and anticipation in sport. In: Williams, A. M. & Jackson, R. C. (Eds.), *Anticipation and decision making in sport.* (pp. 3-24). Routledge: London

Takeuchi, T. & Inomata, K. (2009). Visual search strategies and decision making in baseball batting. *Perceptual and Motor Skills, 108,* 971-980.

Uchida, Y., Kudoh, D., Higuchi, T., Honda, M., & Kanosue, K. (2013). Dynamic visual acuity in baseball players is due to superior tracking abilities. *Medicine & Science in Sports & Exercise, 45*(2), 319-325.

Urgesi, C., Savonitto, M. M., Fabbro, F., & Aglitoti, S. M. (2012). Long- and short-term plastic modeling of action prediction abilities in volleyball. *Psychological Research*, *76*(4), 542-560.

Vater, C., Williams, A. M., & Hossner, E. J. (2020). What do we see out of the corner of our eye? The role of visual pivots and gaze anchors in sport. *International Review of Sport and Exercise Psychology, 13*(1), 81-103.

Ward, P., & Williams, A. M. (2003). Perceptual and cognitive skill development in soccer: the multidimensional nature of expert performance. *Journal of Sport & Exercise Psychology, 25,* 93-111.

Ward, P., Williams, A. M., & Bennett, S. J. (2002). Visual search and biological motion perception. *Research Quarterly for Exercise and Sport, 73*(1), 107-112.

Weissensteiner, J. R., Abernethy, B., Farrow, D. & Muller, S. (2008). The development of anticipation: A cross-sectional examination of the practice experiences contributing to skill in cricket batting. *Journal of Sport & Exercise Psychology, 30,* 663-684.

Williams, A. M. & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. *Research Quarterly for Exercise and Sport, 69*(2), 111-128.

Williams, A. M., Ford, P. R, Eccles, D. W, & Ward, P. (2011). Perceptual-cognitive expertise in sport and its acquisition: Implications for applied cognitive psychology. *Applied Cognitive Psychology, 25*, 432-442.

Wood, J. M. & Abernethy, B. (1997). An assessment of the efficacy of sports vision training programs. *Optometry and Vision Science, 74*(8), 646-659.

Appendix

Table A1: Mean proportional accuracy as a function of player position, pitcher, pitch-type, and occlusion point

Pitcher A									
	FB			**CB**			**CH**		
	133ms	**266ms**	**366ms**	**133ms**	**266ms**	**366ms**	**133ms**	**266ms**	**366ms**
Hitters	0.7	0.57	0.5	0.54	0.59	0.49	0.39	0.48	0.51
Pitchers	0.92	0.71	0.54	0.38	0.42	0.46	0.38	0.17	0.25
Equal	0.58	0.58	0.67	0.5	0.25	0.17	0.25	0.25	0.67
Catchers	0.83	0.75	0.83	0.67	0.42	0.75	0.08	0	0.25
Pitcher B									
	FB			**CB**			**CH**		
	133ms	**266ms**	**366ms**	**133ms**	**266ms**	**366ms**	**133ms**	**266ms**	**366ms**
Hitters	0.81	0.81	0.78	0.78	0.92	0.89	0.23	0.34	0.39
Pitchers	0.83	1	0.96	0.75	0.92	0.96	0.42	0.38	0.46
Equal	0.83	0.75	0.75	0.67	1	1	0.42	0.42	0.5
Catchers	0.5	0.92	0.75	0.42	0.67	0.92	0	0.33	0.42

Table A2: Mean number of hits, false alarms and d prime values as a function of player position, pitcher, pitch-type, and occlusion point

Pitcher A						
	Hit			**False Alarm**		
	133ms	**266ms**	**366ms**	**133ms**	**266ms**	**366ms**
Hitter	0.69	0.66	0.55	0.5	0.58	0.36
Pitcher	0.9	0.7	0.54	0.34	0.7	0.66
Equal	0.58	0.58	0.58	0.35	0.65	0.27
Catcher	0.81	.073	0.73	0.81	0.73	0.73

Pitcher B						
	Hit			**False Alarm**		
	133ms	**266ms**	**366ms**	**133ms**	**266ms**	**366ms**
Hitter	0.8	0.8	0.77	0.26	0.23	0.12
Pitcher	0.82	0.98	0.94	0.14	0.1	0.06
Equal	0.81	0.73	0.73	0.19	0.19	0.04
Catcher	0.5	0.81	0.73	0.5	0.19	0.12

Table A3: Mean percentage confidence for each position and novices for Pitcher A and Pitcher

	Pitcher A	**Pitcher B**
Hitter	72.68	80.92
Pitcher	47.22	68.89
Equal	61.1	76.39
Catcher	55.56	56.67
Novice	52.22	58.06