Journal of Expertise
2020. Vol. 3(2)

ISSN 2573-2773

It Ain't What You Do—It's the Way That You Do It: Is Optimizing Challenge Key in the Development of Super-Elite Batsmen?

Benjamin D. Jones[1], Lew Hardy[1], Gavin Lawrence[1], Ludmila I. Kuncheva[2], Raphael Brandon[3], Mo Bobat[3], and Graham Thorpe[3]
[1]Institute for the Psychology of Elite Performance, Bangor University, Wales, UK
[2]School of Computer Science and Electronic Engineering, Bangor University, Wales, UK
[3]England and Wales Cricket Board, London, UK

Correspondence: Benjamin D. Jones, benjaminjones0411@gmail.com

Abstract

The present study compares the development experiences and the nature and microstructure of practice activities of super-elite and elite cricket batsmen, domains of expertise previously unexplored simultaneously within a truly elite sample. The study modeled the development of super-elite and elite cricket batsmen using non-linear machine learning (pattern recognition) techniques, examining a multitude of variables from across theoretically driven expertise domains. Results revealed a subset of 18 features, from 658 collected, discriminated between super-elite and elite batsmen with excellent classification accuracy (96%). The external validity of this new model is evidenced also by its ability to classify correctly the data obtained from six unseen batsmen with 100% accuracy. Our findings demonstrate that super-elite batsmen undertook a larger volume of skills-based practice that was both more random, and more varied in nature, at age 16. They subsequently adapted to, and transitioned across, the different levels of senior competition quicker. The findings suggest that optimizing challenge at a psychological and technical level is a catalyst for the development of (super-elite) expertise. Application of this holistically driven, non-linear methodological approach to talent pathways and other domains of expertise would likely prove productive.

Keywords

Talent identification, scouting, talent development, coaching, deliberate practice, microstructure of practice, contextual interference, pattern recognition

Introduction

Setting the Scene

Current knowledge from expertise research suggests that the attainment of expertise is highly likely the end-result of an enormously complex interaction between genetic and developmental features (Baker & Cobley, 2013).[1] In a recent review, Rees et al. (2016) argue that differences in early experiences, preferences, opportunities, habits, training, and practice activities are the strongest determinants of mastery in the development of expertise. These differences possess varying importance at different stages of development. Conversion of "giftedness" into "talent" is suggested to result from the accumulation of *desirable* developmental experiences (Gagné, 2004). Therefore, comparing the developmental histories and practice biographies of performers with comparable levels of expertise—and who have maximized their potential—could lead to the identification of the determinants necessary for nurturing expertise.

Deliberate Practice: Sufficient or Necessary for Expertise Attainment?

The strong and positive association between volume of domain-specific practice and the attainment of expertise is grounded in research by Ericsson, Krampe, and Tesch-Römer (1993). The findings highlighted that expert musicians had on average accumulated over 10,000 hours of "deliberate practice" by age 20, while amateurs had accumulated only 2,000 hours, suggesting that deliberate practice is a precursor of mastery. These findings led to the development of the deliberate practice theory, which advocates a mechanism for developing expertise centered on modifying the difficulty of practice commensurate with the skill level of the performer. The theory is centered on the monotonic benefits assumption associated with early specialization, whereby the amount of time engaged in deliberate practice is monotonically related to the individual's acquired performance.

Despite acknowledgement of deliberate practice benefits for the development of sporting expertise, studies examining the average quantity of total practice undertaken by elite sportsmen during development consistently report significant differences to the 10,000 hours over 10 years suggested for musicians (Ericsson et al., 1993; Gladwell, 2008); e.g., cricketers: 7,273 hours (Weissensteiner et al., 2008). Ford et al. (2010) found that practice volume differentiates only high and low performing cricket batsmen between the ages of 13 and 15. Reported differences in practice volume across developmental stages cast doubt over claims that a minimum of 10 years of prolonged practice is required for the attainment of expertise (Ericsson et al., 1993). Moreover, the finding suggests that wider considerations of the microstructure of practice, including practice *type*, *structure*, and *time* when this is carried out, could have more influence on the development of sporting expertise rather than exclusively *how much* practice is accrued. In this regard, emerging research suggests that the reported relationship between age of specialization and practice volume is not necessarily linear, since both elite athletes and cricketers are reported to have undertaken a larger volume of domain-specific practice, compared to the sub-elite, despite specializing later in development (Güllich, 2019; Jones et al., 2019). Moreover, the operationalization of deliberate practice does not account for the potential moderating effect that the microstructure of practice could have on the development of expertise. This presents a barrier to sport officials wishing to structure talent development pathways optimally.

Talent Development

In addition to deliberate practice theory, a number of talent development models originate from the psychology, physiology, education or pedagogy disciplines: Developmental Model of Sports Participation (Côté, Baker, & Abernethy, 2007); Long-Term Athlete Development (Balyi & Hamilton, 2004); Differentiated Model of Giftedness and Talent (Gagné, 2004); Athletic Talent Development Environment model (Henriksen et al., 2010). All these models have advanced our understanding of expertise development and filled a gap between theory and applied practice. That said, their generic nature presents challenges for identifying "optimal" practice environments in sport (see Phillips, Davids, Renshaw, & Portus, 2010). These challenges could partly be attributed to the additive effects observed within most talent development models suggested to develop exceptional performance. However, the influence of microstructure of practice on *how much* practice is necessary for developing expertise in sport remains to be explored.

The current literature is also limited by a lack of understanding of the interactions taking place between the nature and microstructure of practice activities and wider developmental histories to develop expertise. Weissensteiner et al. (2008) explored the features of developmental history that contribute to the acquisition of skilled cricket batsmen. Using discriminant function analyses, they aimed to determine features which most accurately discriminated between high or low-performing batsmen (categorized according to anticipation ability). The study highlighted that accrued practice volume was a weak predictor of anticipatory skill. The authors suggested that their measures of practice experience adopted may have been insufficiently fine grained, lacking the sensitivity required to capture the critical elements of practice experience that contribute to the acquisition of anticipatory skill. Furthermore, the study explains how acquisition could be more closely related to the type of cricket-specific practice undertaken, rather than the quantity, thus highlighting a need to precisely measure the microstructure of practice. All said, there is a limited body of research that has examined the microstructure of sport practice within the expertise development field.

An Introduction to Contextual Interference

Much of the motor learning research pertaining to the microstructure of practice has emanated from controlled laboratory experiments with unskilled participants and over short learning periods. In this setting, the contextual interference effect on practice has been most widely researched (for a review, see Brady, 2008). The contextual interference effect stipulates that multiple skills (or skill variations) are more effectively learned when there is interference present during practice (for a review, see Monsell, 2003). At a basic level, the interference can be created by manipulating the structure of practice trials such that skills are learned in either a blocked or random fashion.

Random scheduling enforces the learner to switch between the skills "randomly" throughout practice, whereas blocked practice requires the learner to practice one skill for a block of repetitions before switching to the other skill (Farrow & Buszard, 2017). One likely conclusion is that although random practice has detrimental effects on performance during short-term acquisition, it facilitates learning in the long term. This is achieved either by encouraging the performer to undertake more elaborate and distinctive processing from one trial to the next (i.e., the elaboration hypothesis; Shea & Morgan, 1979) or through forgetting and subsequently reconstructing an action plan each time that a skill is performed (i.e., the action plan reconstruction hypothesis; Lee & Magill, 1985). The benefits of contextual interference extend to skills which demand the same class of actions (e.g., executing different cricket batting shots) through practicing different variations of these skills (e.g., manipulating the direction, loft, pace of a batting shot), known as variable practice (Schmidt & Bjork, 1992). This is the opposite to constant practice, where the parameters of a skill are instead fixed. Indeed, the benefits of variable practice are greatest when schedules of practice are somewhat unpredictable (Porter & Magill, 2010). Despite the environmental constraints of this research, random practice, combined with variable practice, could result in superior long-term skill retention, especially for performance scenarios which are somewhat unpredictable, and demand both the rapid retrieval of movement skills and extreme accuracy in their execution (i.e., typical characteristics of expert performers) (for a review, see Monsell, 2003). Thus, prolonged random and variable practice could conceivably aid the development of cricket batsmen, specifically through challenging players to develop and execute run scoring based on situational information.

Random and Variable Practice: A Mechanism for Optimizing Challenge?

Experimental research has demonstrated that high contextual interference places exceedingly high demands on cognitive processing (Broadbent et al., 2017), which could potentially inhibit the benefits typically found to emerge from such practice in laboratory settings. Hence, task difficulty, or skill complexity, relative to the performer, appear central factors in moderating the contextual interference effect. This position is consistent with the various accounts of learning, whereby learning is more robust when the task difficulty presents an optimal challenge to the performer (e.g., Challenge Point Framework, Guadagnoli & Lee, 2004; Deliberate Practice, Ericsson et al., 1993).

Despite receiving extensive coverage, the scheduling of practice represents a single method, from potentially many viable methods, for increasing task difficulty via contextual interference. The method and precision by which performers strategize and execute their actions is individualized and heavily influenced by level of expertise (Gentile, 1972; Khan et al., 2006). This raises a question in the case of youth performers, who are typically arranged into age group bands, and could receive exposure to similar practice structures as a result: *To what extent is it possible for group practice be optimized at an individual level - despite differences in performers' stage of development?* The contextual interference effect denotes that practices structured to contain interference facilitates long-term skill retention, despite likely being detrimental to performance in the short-term (Porter & Magill, 2010; Shea & Morgan, 1979). That said, performers who are more advanced in their development, relative to their peer group, may be less challenged by general group practices, and face less performance detriments in practice as a result (Guadagnoli & Lee, 2004). In this regard, the higher skilled performers in a group, who are less occupied by their potential technical inadequacies (Gentile, 1972), can conceivably divert more attention to planning and execution strategies within group practice settings, increasing practice interference (over and above that posed by the structure of group practice), to achieve *individualized* practice outcomes.

While knowledge pertaining to the nature and microstructure of practice largely stem from lab-based research with novices, we can reasonably theorize that these learning domains extend to the development of expertise in sport. That said, while there is clearly a place for such lab research, the literature is at a point where there is a need to validate the findings in the field (Farrow & Buszard, 2017). In this regard, combining random and variable practice, and gradually increasing contextual interference, as a function of task difficulty and skill complexity relative to the performer's stage of development, could serve as a function for optimizing challenge for cricket batsmen (Guadagnoli & Lee, 2004).

The Specificity of Practice Principle

The superior learning associated with random and variable practice conditions likely reflect the benefits of representative learning/practice design (Pinder et al., 2011). This extends the specificity of practice principle, which denotes that practice conditions closely matching the movements of the target skill and the conditions of the target context, result in optimal learning (Henry, 1968; Rothwell et al., 2017). In sport, competition constitutes the target context, and competitive performance represents the intended output of learning. Random and variable practice could be particularly beneficial in an open loop sport, such as cricket, where a batsman's output is in direct response to the somewhat unpredictable opposition bowlers' deliveries (Porter & Magill, 2010). More so, when considering that the demands of international cricket require batsmen to adapt, often required to produce multiple shot types in succession, in response to bowler deliveries, and manipulate the direction, loft, and pace of shots (variability) in response to this and wider situational information. A problem associated with the traditional scheduling of practice is the development of skills in a non-pressurised environment as a pre-requisite for performance of skills in pressurised situations, whereas competition demands the production of skills under pressure (Lawrence et al., 2014). Therefore, prolonged exposure to the inherent challenge of practice conditions which closely matches the movements of competition, during early development, could facilitate effective and consistent skill execution in pressurized conditions.

Summary of Limitations

In summary, the current literature provides limited understanding of the interaction between developmental characteristics *and* the microstructure of practice activity. The current divide between research and applied sport practice is highlighted by the absence of research examining the nature and microstructure of practice in elite performers. This imbalance exists despite being widely believed as important, both from a theoretical *and* an applied perspective. Consequently, if future research is to achieve a better understanding of optimal development environments, sport-specific examinations of the nature and microstructure of practice activity, alongside developmental experiences, are warranted to identify the following: the skills practiced, how practice is structured and delivered, how frequently this is practiced, and how this practice changes over the course of development. Pattern recognition analysis offers promise in addressing these questions, given its ability to model the multiple and complex interactions between multidisciplinary features (variables), and accounting for the multifaceted and dynamic nature of expertise, and reflecting a holistic approach to identifying precursors of expertise. This methodology was recently applied to identify predictive features that discriminate between samples of elite and sub-elite cricket spin bowlers (Jones et al., 2019) and super-elite and elite Olympians (Güllich et al., 2019).

Study Rationale

The present study is the first known to have applied a framework to measure the contextual interference and variability of practice effects among a truly elite sample. Furthermore, the study comprehensively explores the multifaceted and dynamic nature of expertise by examining the nature and microstructure of cricket batting practice against the developmental histories of super-elite batsmen using advanced non-linear pattern recognition techniques. This approach overcomes existing limitations in also allowing for a more fine-grained approach to exploring the influence of the microstructure of practice on the practice volume-development of expertise relationship. The approach was expected to identify cricket and batting-specific precursors of expertise, most predictive of elite performance. The research findings will enable a greater overall understanding of the relative importance of batsmen's development provisions and experiences, while also leading to a greater understanding of the specific interactive features common to super-elite batsmen that contribute to their holistic development, and enable the benchmarking of these precursors.

Method

Participants

The total sample comprised 20 past and present batsmen, 10 of whom were super-elite (M_{age} = 36; *SD* = 6.3) and 10 elite (M_{age} = 34; *SD* = 3.6). Super-elite batsmen were sampled on the basis of the following three criteria, and were applied in order of appearance: (1) had played for the England national team post-2004 (M_{innings} = 247; *SD* = 67); (2) possessed a robust technique that enabled them to thrive against world class pace or spin bowling; (3) continuously produced match-winning performances for England in Test or limited overs formats “when it mattered”.[2] Elite batsmen were sampled on the basis that they had maintained prolonged careers at the highest standard of domestic cricket, by playing in a minimum of 100 innings of First-Class County Cricket innings (M_{innings} = 279; *SD* = 110), and represented the pool from which all super-elite batsmen had emerged. However, none of the elite batsmen had played for England in any senior competition; elite batsmen still playing were deemed unlikely, by the England & Wales Cricket Board’s (ECB) National Lead Batting Coach, to represent England in the future, owing to their age. The elite batsmen selected for the study were subsequently matched to the individual super-elite batsmen based on three characteristics: (1) career era (*played First-Class County Cricket post-2004*); (2) batting position (*opening/top order/middle order*); (3) educational background (*public/state schooling*). A clear distinction exists in the performance levels reached by the elite and super-elite; the super-elite represent a subsample of just 2% of English batsmen who played First-Class County Cricket within the same era (2004-2016). This clear distinction in participants' level of expertise allowed a robust examination of the precursors of super-elite expertise (Table 1).

Table 1. Overview of the theoretical domains explored with the "Attainment of Batting Expertise Interview Schedule"

	Super-Elite	Elite
Common Criteria	Sample comprised England-qualified elite and super-elite cricket batsmen All participants initially sampled from First Class County cricket Participants grouped into pairs and matched on career era, batting position, and educational background	
Age	36 Years ± 6.3	34 Years ± 3.6
# of Competition Innings Played	First Class County: 407 ± 158 England senior team: 247 ± 67	First Class County 270 ± 110
Unique Criteria	1) Represented England senior team post-2004 2) Possess(ed) a technique to thrive against world class pace/spin 3) Continuously produced match-winning performance for England "when it mattered"	1) Maintained prolonged careers in First Class County Cricket *(Min. 100 innings)* 2) Had never represented England at senior level/or were deemed unlikely to

Table 2. Overview of the theoretical domains explored within the "Attainment of Batting Expertise" interview schedule

Structured Batting Interview	
Section 1: Demographic Information • Birthdate • Birthplace • Homeplace • Parental sporting history and achievement • Parental coaching experience • Sibling order effect • Schooling type and experiences • Academic achievements and milestones	**Section 2: Developmental Sporting Activity** • Volume of cricket activity (play, practice, and competition • Number and type of general sports • Prevalence of deliberate play and deliberate practice • Sport and cricket ages (accumulated experience) • Early cricket specialization vs. sport diversification • Batting specialization age • Linearity of development in cricket (academy/county teams inclusion and exclusion frequency)
Section 3: Developmental Milestones and Performance Indicators • Highest level of cricket representation by ages 16, 18, and 22 • Age selected for all representation levels • Level of technical and psychological challenge • Time taken to achieve significant performances • Age became team's best/one of best batsmen • Perceived quality of coaching and facilities • Injury time across defined time periods	**Section 4: Nature and Microstructure of Practice** • Deliberate play and deliberate cricket activity • Physical fitness activity • Mental skills training • Vicarious learning • Conveyance of instruction • Batting practice structure and bowling delivery types and methods faced • Decision making/execution difficulty • Context and anxiety specificity • Internal and external foci of attention (and nature) • Intrinsic and extrinsic feedback • Constraints and prescriptive coaching approaches

Note: N = 658 quantitative features were collected from the interview for each participant.

Measures

Attainment of Batting Expertise Interview Schedule. A structured interview schedule was developed comprising four sections (Table 2). Section 1 (Demographic information), section 2 (Developmental sporting activity), and section 3 (Cricket developmental milestones & performance indicators) of the interview schedule were informed by previous research exploring precursors of expertise (Côté. Ericsson, & Law, 2005; Hardy et al., 2013; Jones et al., 2019). These sections encompassed questions surrounding batsmen's development from the age of 6 to 22. The existing interview schedules were refined to achieve a better understanding of optimal development environments in cricket. Section 4 was developed specifically by the researchers for the present study to address the dearth of research exploring the influence of the microstructure of practice on the development of sporting expertise. This section measured the microstructure of the batsmen's practice at key developmental stages of the ECB player pathway (ages 16, 18, and 22), as identified by the ECB's Head of Science, Medicine, and Innovation. The questions in section 4 centered on the specific time-point of the cricket calendar that participants had reported engaging in the largest volume of practice (summer or winter). It was hoped that this method would alleviate some of the well-documented limitations with regards to retrospective recall, specifically surrounding the accuracy of responses provided (e.g., Hopwood, 2013), by focusing on the time-point that each participant recalled doing most practice in at each specified age. The developed interview schedule was then subjected to a three-stage piloting process. First, the ECB's Head of Science, Medicine, and Innovation reviewed the interview schedule and provided detailed constructive feedback for refinement. Second, the schedule was piloted on a number of elite batsmen and England Development Program batting coaches to assess the relevance of theoretical content against the structure and terminology of the player pathway. A final pilot interview was then performed with the Director of England Cricket, who subsequently approved the study. The final interview schedule (comprising four expertise domains) can be found in the Supplementary Information.

Methodological Design

Super-elite sportsmen are, by definition, extraordinary, and we adopted multi-level, stringent criteria to represent their superior level of expertise, a sample classification method advocated by Jones et al. (2018). Consequently, the present study addressed inconsistencies observed in the sampling classification methods of previous research that were due to simplistic dichotomization of level of expertise (Coutinho et al., 2016). The batsmen's existing level of expertise demonstrate that, overall, the effects of their developmental experiences and practice histories are durable, meaning that identifying the enduring discriminating factors will go some way toward addressing the drawbacks of short transfer effects in previous research. The super-elite sample was identified first, and the elite participants were subsequently matched (*career era, batting position,* and *educational background*) according to a matched-pair design, a design similar to that used in Hardy et al.'s (2013) seminal study. Matching participants on the identified key characteristics assisted in exploring why batsmen digress in their eventual expertise despite their common characteristics, thus enabling the present study to address the "what makes the difference?" question comprehensively across the four interview sections outlined. The quantitative dataset comprised 20 participants (objects), with 658 features (variables), and this self-reported data was directly put into MS Excel during the interviews and collated prior to analysis.

Procedure

Following institutional ethical approval for research involving human participants, the participants were recruited by the Director of England Cricket and the National Lead Batting Coach. All participants provided written informed consent in advance of interview. Each structured interview lasted approximately 3 hours, was recorded using a digital Dictaphone, and was designed such that all data collected were quantitative. Once all interviews had been completed, data were standardized, and then analyzed using pattern recognition approaches, with the primary aim of determining the features from the practice biographies and developmental histories of batsmen that best discriminate between the super-elite and elite.

Analytical Strategy Overview

Pattern recognition analysis has been developed in bioinformatics to solve the problem of classifying objects based upon their features (Hastie et al., 2003), and has recently been applied within sport sciences. The analysis offers a non-linear approach to analyze multidimensional data to represent the multifaceted and dynamic nature of expertise.

Pattern recognition analysis overcomes the limitations of linear techniques, which typically either combine features (variables) additively or analyze features in isolation (Jones et al., 2019). This method employs modern computational power to analyze iteratively a large number of features in order to identify the pattern of features that most accurately discriminate between different classes of objects (participants). Pattern recognition typically comprises 3 stages: feature selection, classification, and recursive feature elimination (for a detailed description of these procedures, *see* Güllich et al., 2019).

Feature selection identifies the individual predictive features that best discriminate between (the super-elite and elite) classes. Pattern recognition analysis requires a robust method of feature selection for such a "wide" data set where there are far more features than objects. The four feature selection methods utilized in the present study have been chosen because of their suitability for use with wide data-sets: Support Vector Machine (SVM; Burges, 1998); Relief-F (Kira & Rendall, 1992); Fast Correlation Based Filter (FCBF; Yu & Liu, 2003); and Correlation Attribute Evaluation (Hall, 1999). These four feature selection methods use very different criteria, consequently, the more times that a common feature is selected by different feature selection methods, the greater confidence can be placed in that feature's predictive power, preventing spurious results.

Classification involves the analysis of a specified subset of features, with the aim of discriminating between groups of classes. In the present study, feature subsets are derived from the feature selection protocol; the super-elite and elite represent the predefined classes. Thus classification accuracy is determined by the number of batsmen that are correctly assigned as super-elite or elite. Once again, greater confidence can be placed in feature sets that have consistent rates of classification accuracy. Consequently, four different classifiers were applied to the feature subsets selected in the present study: SVM (as used in the feature selection; Burges, 1998); Multilayer Perceptron (MLP; Bishop, 1995); Naïve Bayes (NB; Hand & Yu, 2001); and Nearest Neighbor (Lazy learner, IB1; Duda et al., 2001).

Recursive Feature Elimination (RFE) (Guyon et al., 2002), also known as "fitting," is a procedure that identifies the subset of features that predicts the class labels with higher classification accuracy, thus allowing us to provide the user with the optimal solution for a given data-set (Güllich et al., 2019; Jones et al., 2019). RFE is applied to subsets usually consisting of a large number of features where fewer, as opposed to greater, features are likely to offer the optimal solution.

Analytical Strategy Summary

In the present study, the predictive power of the 658 features collected was assessed by ascertaining how accurately they discriminated between the super-elite and elite batsmen. In order to extract discriminatory features from the data, we used the Waikato Environment for Knowledge Analysis (Weka; Hall et al., 2009). Weka is a machine learning workbench that offers a wide range of algorithms for data pre-processing, feature selection, and classification. Feature selection and classification methods were subjected to leave-one-out cross-validation (LOO), to mitigate the risk of overfitting and to provide a more realistic prediction of the classification function on unseen data (generalization performance) (Kuncheva & Rodríguez, 2018). The analytical strategy adopted in the present study is based on the strategy of Güllich et al. (2019).

Section Analysis. The first stage of the analysis involved applying the *feature selection* protocol to identify separately the predictive power of features from each of the four expertise domains of the interview schedule (demographic information, developmental sporting activity, developmental milestones and performance indicators, and the nature and microstructure of practice). Features from each section possessing the greatest predictive power were subsequently pooled together; the predictive power of features was determined by the consistency with which they appeared in the top-20 features selected by each of the outlined four feature selection methods. Using this procedure, three subsets of predictive features were selected, according to three different degrees of stringency (A, B, C) (see Figure 1, next page):

- **Feature Subset (A):** Features ranked in the top 20 discriminatory features by at least two out of four feature selection methods *(least rigorous/most liberal).*
- **Features Subset (B):** Features ranked in the top 20 discriminatory features by at least three out of four feature selection methods.
- **Features Subset (C):** Features ranked in the top 20 discriminatory features by all four feature selection methods *(most rigorous/most conservative).*

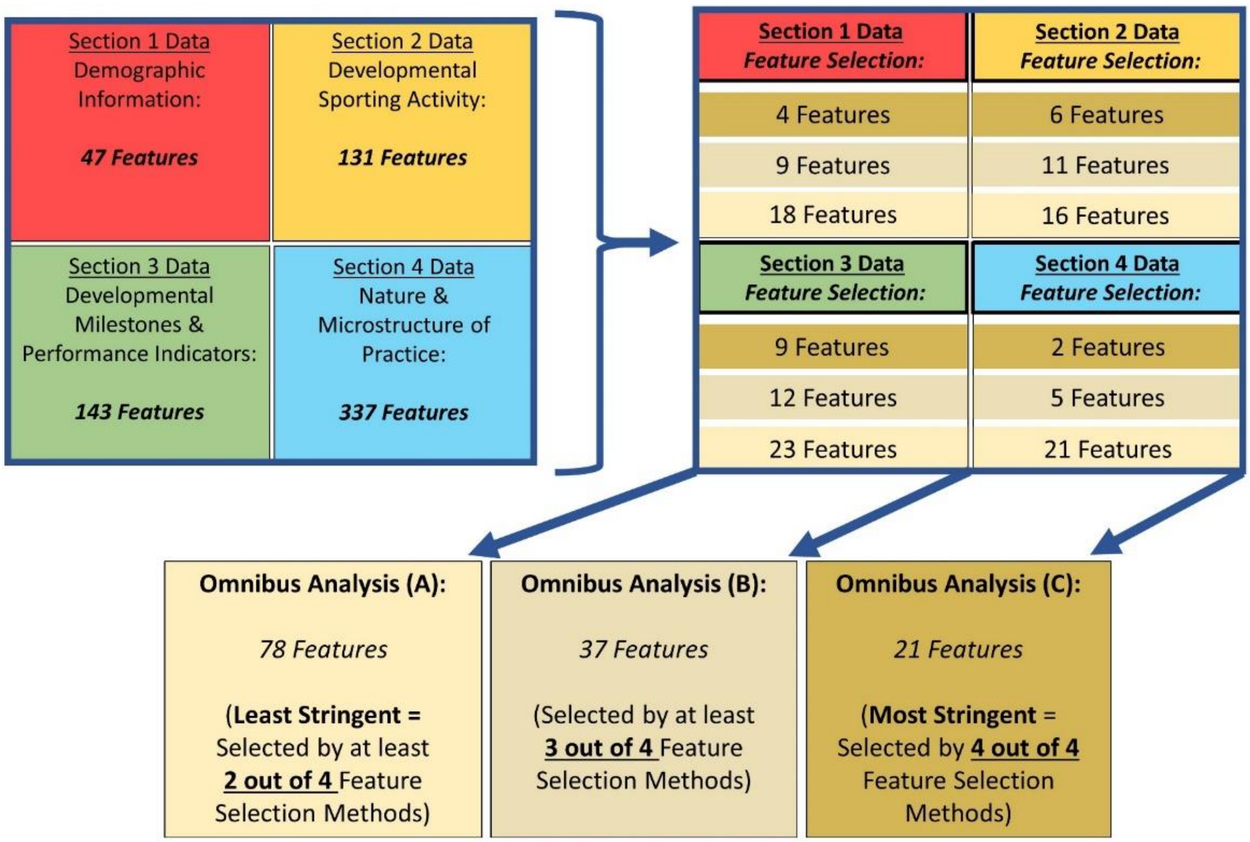

Figure 1. Feature selection summary for the section analysis: The consistency by which features appeared within the top-20 features for each of the four feature selection methods, creating three subsets of features with different degrees of stringency

Subset Analysis. The subsets from each of the four expertise domains were then combined to perform a set of three omnibus analyses with varying degrees of stringency (A, B, C); these subsets cumulatively totalled 78 features. Considering this substantial number of features that existed across the subsets, the first step of the omnibus analysis involved repeating the feature selection procedure within each subset, to assess the *relative* predictive power of their amalgamated features. Following this, classification protocols were applied, using the four classifiers outlined, to assess the combined discriminative power of the three feature subsets produced. For each of these subsets (A, B, C), the feature subset producing the highest overall classification accuracy was selected and is presented in Table 3 (page 154). Recursive Feature Elimination method (RFE) was subsequently applied to two of the three feature subsets selected to arrive at an "optimal" solution in the case of each subset by only retaining the fewest number of features that discriminate between classes with the greatest accuracy. Finally, the three reduced (optimal) solutions were amalgamated into a single, final classification analysis, and are reported in the results section.

Results

Final Classification Model: Overview

The omnibus analyses produced three different solutions (A, B, C), each discriminating with excellent accuracy between the super-elite and elite batsmen. Each solution reflects the result of slightly different feature selection, classification, and recursive feature elimination conducted during the omnibus analyses (see Figure 1). These three solutions collectively contain a total 18 different features (which do not all appear in any one solution), and, for the sake of inclusiveness, the 18 features were put into a combined final classification model, also producing excellent accuracy (M = 96.25%). The accuracy of each classifier is listed below:

- Support Vector Machine (SVM) Classifier: 100%
- Multilayer Perceptron (MLP) Classifier: 100%
- Naïve Bayes Classifier: 90%
- Nearest Neighbor (Lazy learner, IB1) Classifier: 95%

Final Classification Model: Summary

Descriptive statistics and the direction of the 18 discriminating features are presented in Table 4 (page 155). Results from the comparison of super-elite and elite batsmen demonstrate that the super-elite have these characteristics:

1. Have *more* siblings who are older
2. Engaged in a *larger* volume of cricket practice activity at age 16
3. Undertook a *larger* volume of cricket practice within their busiest practice period at age 16
4. Were engaged in a *larger* volume of cricket play at age 16
5. Practiced a *greater* number of shots during their random batting practice at age 16
6. Undertook a *larger* volume of random-variable batting practice with maximum variation (3 variations) at age 16
7. Took *fewer* years to transition between the highest level of club cricket played by age 16 to their First XI County Cricket Debut
8. Became the best batsman in their Second XI County Cricket team at a *younger* age
9. Made their List A (professional) cricket debut at a *younger* age
10. Were *older* when selected for their highest level of general cricket competition played by the age of 18[3]
11. Missed *less* development time through injury between ages 19 and 22
12. Were *younger* when selected for their highest level of county cricket competition played by age 22
13. Experienced a *larger* volume of cricket competition at age 21
14. Accumulated a *larger* volume of total cricket activity at age 21
15. Experienced a *larger* volume of cricket competition at age 22
16. Became one of the best batsmen in their First XI County Cricket team at a *younger* age
17. Were *more likely* to become the best batsman in their First XI County Cricket team
18. Became the best batsman in their First XI County Cricket team at a *younger* age

The clear distinction in the 18-feature holistic development profiles of the super-elite and elite are presented in Figure 2 (page 156) and are depicted on a developmental timeline in Figure 3 (page 156).

The multistage approach of the analyses is underpinned by the premise that the more times a common feature appears across the different solutions, the more confidence that can be placed in the feature's importance. This consensus is displayed in Table 5. The table highlights that 6 features, from a possible 18, were contained in all 3 solutions, demonstrating high consistency. An additional 3 features were contained in 2 of the 3 solutions, demonstrating moderate consistency. The remaining 9 features were contained in 1 of the 3 solutions, demonstrating relatively low consistency (but high accuracy; see Discussion for implications). An important disclaimer must be made here. The classification accuracies which we report for the above analyses may be slightly optimistically biased. The reason is because Weka's protocol for feature selection (LOO cross-validation or not) is followed by another round of using the same data in order to train and test the classifier (LOO). In other words, the object set aside for testing has been "seen" during the previous training-and-testing protocol when feature selection was carried out. That said, this so-called "peeking" (Kuncheva, 2014) effect is indirect and ignored in many studies. Nonetheless, one cannot make the claim that the classification accuracy on unseen data would exactly match the one achieved for this dataset, until the model has been directly tested (performed as part of "Confirmatory Model Testing" below).

Table 3. Summary of the best solutions produced from the omnibus analyses

	Omnibus A	**Omnibus B**	**Omnibus C**
Features Put In	78	37	21
Number of Features Selected in Best Solution	19	9	17
Initial Classification Accuracy (Average)	92.5%	98.75%	91.25%
Number of Features Omitted	5	0	7
Final Solution: Number of Features	14	9	10
Final Solution: Classification Accuracy (Average)	98.75%	98.75%	98.75%
Final Solution: Feature Descriptors	- Volume of cricket play age 16 - Volume of cricket practice activity within busiest practice period age 16 - Volume of random-variable batting practice with maximum variation (3 variations) - Age selected for highest level of cricket competition by age 18 - Age selected for highest level of county cricket by age 22 - Age made senior list a (professional) debut - Age became the best batsman in their second XI county cricket team - Development time missed through injury between ages 19-22 (months) - Volume of cricket competition age 21 - Volume of total cricket activity age 21 (practice + competition) - Volume of cricket competition age 22 - Age became one of the best batsmen in their first XI county team - Became the best batsman in their first xi county team (outright) - Age became the best batsman in their first XI county team	- Volume of random-variable batting practice with maximum variation (3 variations) - Age selected for highest level of cricket competition by age 18 - Age made senior list a (professional) debut - Age became the best batsman in their second XI county team - Volume of cricket competition age 21 - Volume of total cricket activity age 21 (practice + competition) - Volume of cricket competition age 22 - Became the best batsman in their first XI county team (outright) - Age became the best batsman in their first XI county team	- Number of older siblings - Volume of cricket practice activity age 16 - Number of shots practiced randomly age 16 - Volume of random-variable batting practice with maximum variation (3 variations) - Years to transition from club cricket at age 16 to first XI county cricket team - Age selected for highest level of cricket competition by age 18 - Age became the best batsman in their second XI county team - Volume of cricket competition age 21 - Volume of total cricket activity age 21 (practice + competition) - Age became the best batsman in their first XI county team

Table 4. Unstandardized descriptive statistics of the 18 features of development that discriminate between super-elite and elite batsmen.

#	Feature	Direction *(+ / -)*	Super-elite			Elite		
			Mean	**Median**	***SD***	**Mean**	**Median**	***SD***
1	Number of older siblings	+	1.20	1.00	1.07	.40	0	.91
2	Volume of cricket practice activity age 16	+	355.00	401.70	167.00	198.00	201.50	36
3	Number of shots practiced randomly age 16	+	10.20	11.00	2.00	8.00	9.00	1.94
4	Volume of cricket play age 16	+	129.72	102.29	86.09	42.69	22.37	38.30
5	Volume of cricket practice activity within busiest practice period age 16	+	243.00	260.00	112.00	154.00	138.00	31.00
6	Volume of random-variable batting practice with maximum variation (3 variations)	+	103.35	78.32	79.47	19.50	0	34.88
7	Years to transition from club cricket at age 16 to first XI county cricket team	-	3.40	3.00	1.01	5.40	5.50	2.29
8	Age selected for highest level of cricket competition by age 18	+	17.50	18.00	.67	16.60	16.00	.80
9	Age selected for highest level of county cricket by age 22	-	17.90	18.00	1.04	19.90	22.50	1.92
10	Age made senior list a (professional) debut	-	18.48	18.79	1.04	21.17	21.41	1.94
11	Age became the best batsman in their second XI county team	-	19.50	18.50	2.59	23.30	23.00	2.38
12	Development time missed through injury between ages 19-22 (months)	-	0	0	.20	1.32	.13	1.97
13	Volume of cricket competition age 21	+	867.00	860.00	120.00	528.00	563.01	231.00
14	Volume of total cricket activity age 21 (practice + competition)	+	1206.00	1176.50	158.00	741.00	859.72	299.00
15	Volume of cricket competition age 22	+	865.00	913.99	282.00	526.00	562.75	221.00
16	Age became one of the best batsmen in their first XI county team	-	20.60	20.75	2.24	25.00	26.00	2.87
17	Became the best batsman in their first XI county team (outright)	+	1.00	1.00	.16	.50	.50	.50
18	Age became the best batsman in their first XI county team	-	23.70	23.50	3.20	29.55	30.50	2.23

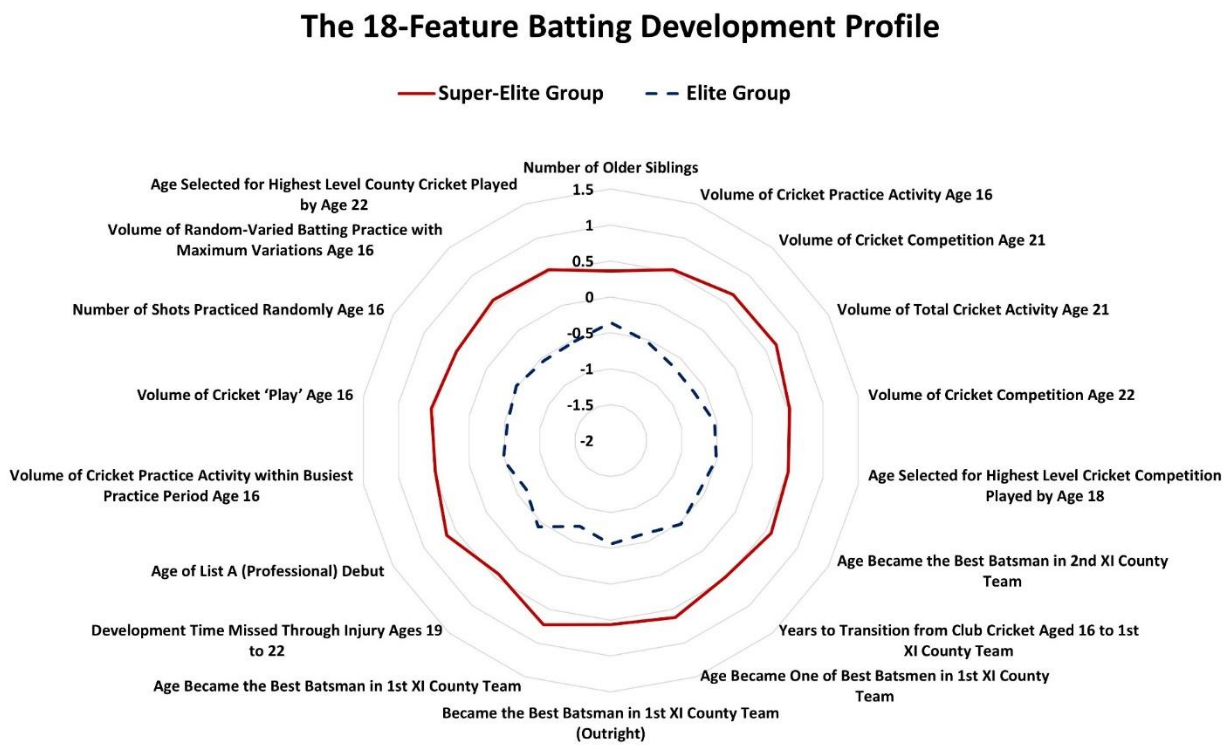

Figure 2. The discriminating development profiles of the super-elite and elite batsmen. *Note:* Data points reflect the standardized mean values for each expertise class. A higher number is associated with the super-elite class. The values of negatively weighted features (outlined in Table 2) are reversed in order to present the discrimination of the super-elite/elite development profiles through visual means.

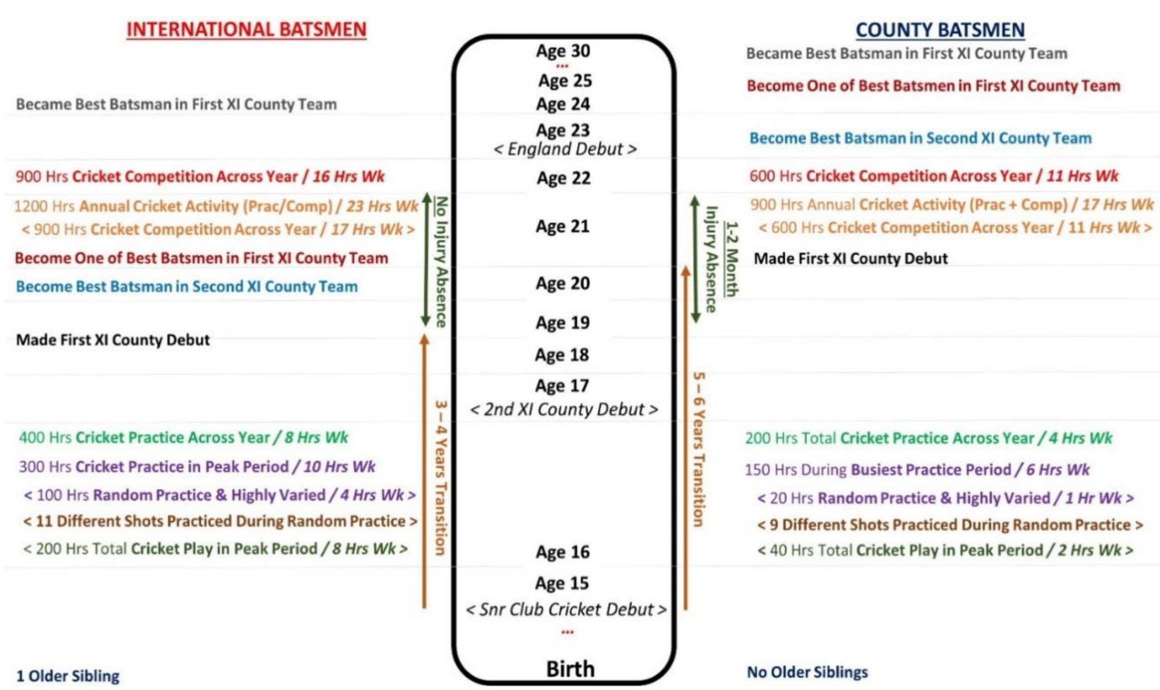

Figure 3. A timeline of the 18 developmental discriminating features between super-elite (left) and elite (right) batsmen. *Note:* Data points reflect the unstandardized median values of each feature (approximation).

Confirmatory Model Testing

The 18-feature model discriminates between the super-elite and elite batsmen with excellent accuracy. The next step was to test this (trained) classification model's ability to generalize (and thus predict) unseen data-sets, i.e., batsmen who were not included in the original analysis. To do this, we utilized the interview data of 6 additional English batsmen, 3 of whom were classified as super-elite, and 3 of whom as elite. The same 4 classifiers ever-present during the omnibus analyses were adopted for model, and the results are reported below:

- Support Vector Machine (SVM) Classifier: 100%
- Multilayer Perceptron (MLP) Classifier: 100%
- Naïve Bayes Classifier: 100%
- Nearest Neighbour (Lazy learner, IB1) Classifier: 100%

Confirmatory model testing revealed 100% classification accuracy across the 4 classifiers, validating the 18-feature-model's generalizability on 6 unseen data sets.

Discussion

The present study developed and employed a novel method to examine the combined contribution of the nature and microstructure of practice, with developmental experiences, to understand "what makes the difference" in the development of super-elite expertise. Results revealed a predictive model containing 18 features, from a possible 658, that discriminated between the super-elite and elite batsmen with excellent accuracy (96%). Subsequent validation analysis of the final 18-feature model which contained an unseen data set of six batsmen, revealed a perfect (100%) classification fit of this testing data across four classifiers used, thus providing early evidence of the model's external validity. Furthermore, the multistage omnibus analyses contained degrees of stringency, enabling different confidence levels to be attached to subsets of the 18 features. The study adds to the extant literature in several ways. First, it examined the microstructure of practice in a sample of truly elite sportsmen and was thereby not restricted to solely "counting hours." Second, it utilized a serial framework that connected theoretical constructs, previously typically examined disparately. Third, the non-linear capabilities of machine learning enabled exploration of the multiple and complex interactions between individual features, thereby contributing a holistic understanding of the multifaceted and dynamic nature of expertise. The discussion follows the temporal sequence of development; the 18 features are subdivided into 3 areas of development: Type and Volume of Activity, Transition, and Adaptability.

Type and Volume of Development Activity

Super-elite batsmen undertook a larger volume of cricket practice at age 16, compared to the elite, across both the calendar year and during their most concentrated period of practice (summer *or* winter). This finding is consistent with the corpus of research attributing the development of expertise to vast quantities of domain-specific practice (e.g., Ericsson et al., 1993; Jones et al., 2019).

Examination of the microstructure of practice at age 16 revealed that the super-elite players had also undertaken a larger volume of random practice with greater variability, discriminating them from the elite. Specifically, the super-elite batsmen reported undertaking a larger volume of practice indicative of "scoring based scenarios," which challenge players to develop and execute run scoring based on game information. In addition, the super-elite's random practice was also more random in nature at age 16, as they practiced a greater number of shots. The volume and type of bowling deliveries that batsmen faced during practice did *not* discriminate, representing a commonality between the super-elite and elite, and therefore indicating that it is the batsmen's output which distinguishes at the super-elite level in the present study (for more information on the types of bowling deliveries measured, see Supplementary Information).

Overall, the findings demonstrate that random practice *and* variability in practice relatively early in batsmen's development (at age 16) are precursors of super-elite expertise. This furthers our conceptual understanding, given that these concepts have typically been researched in *isolation*, and have not previously been concurrently measured in a truly elite sample within an applied setting (Farrow & Buszard, 2017). Although highly random and varied practice is often considered detrimental to performance during early skill acquisition, due to the increased challenge associated with its dynamic nature (Lin et al., 2008), the present findings reaffirms previous evidence of its long-term benefits. One likely explanation for the present findings relates to the superior long-term learning retention associated with higher contextual interference (for a review, see Monsell, 2003). Moreover, by addressing the questions of "what, how and when" one practices—rather than the historically answered question of "how much"—these findings offer a serial framework by which domain-specific practice hours may be constructed within an elite sporting environment,

The mechanism through which the super-elite may develop from performing (challenging) practice poorly during skill acquisition to achieving mastery is intriguing, as it highlights a disparity between the indicators of elite performance at senior and youth levels. Gradual improvement of performance is suggested to be contingent on three conditions: level of challenge, availability of feedback, and opportunity for error detection and correction (Ericsson et al., 1993; Guadagnoli & Lee, 2004). While "optimal" challenge was not directly measured in the present study, the additional information presented by the super-elite's higher volume of more random and varied practice at age 16 is indicative of greater nominal difficulty (challenge), when compared to practice conditions with lower contextual interference and variability (i.e., blocked and constant practice [Shea & Morgan, 1979]). Furthermore, ratings of mental effort and execution difficulty during practice did not discriminate between the super-elite/elite at age 16. This likely represents the functional task difficulty posed by the differing practice conditions relative to each group. Consequently, the present finding suggests that the super-elite's higher contextual interference and variability during their cricket batting practice at age 16 could have been a mechanism for optimizing challenge during learning. This practice, while more challenging, is dynamic and less repetitive; this is demonstrated by the super-elite's reporting that a greater volume of their cricket activity was representative of play than the elite, at age 16 (i.e., fun, free from specific focus, and providing immediate gratification).

Despite the noted function of contextual interference in optimizing challenge within the present study, the mechanism by which the super-elite were exposed to greater contextual interference is less clear. On the one hand, the super-elite's larger volume of random and variable practice could conceivably have been the result of greater exposure to practice environments invoking random and variable practice; i.e., through the scheduling of practice, as theorized in the literature (for a review, see Brady, 2008; Monsell, 2003). However, as it is the batsmen's practice *output* which discriminates the super-elite's practice exclusively (and not the volume and types of bowling deliveries faced), the super-elite's larger volume of random and variable practice could instead reflect their advanced stage of development by the age of 16.

In essence, the greater time spent in highly randomized practice environments may be a reflection of the super-elite's prolonged competitive state of mind during practice, coupled with their added ability to strategize within their practice accordingly, owing to their advanced stage of development by age 16; the super-elite's prolonged specificity within their practice could have facilitated the long-term successful replication of these skills to higher-level competition environments from an earlier age (Henry, 1968; Rothwell et al., 2017).

Super-elite batsmen have more older siblings than elite batsmen; this is consistent with past research at the elite level, where having an older sibling is a common circumstance in performers (Hopwood et al., 2015). The present finding represents a pronounced sibling effect; we suggest that this finding reflects heightened competitive exposure to multiple older siblings. These challenging sibling dynamics can foster resilience and equip performers for coping with future high-level challenges (MacNamara et al., 2010).

The super-elite's greater competition volume at ages 21 and 22 discriminated them from the elite; this period represents the two years preceding their international debut (M_{age} = 23). The super-elite's greater cricket activity volume (practice + competition) at age 21 is a product of their larger competition volume at that age. These findings are consistent with research demonstrating that elite (international) cricket spin bowlers experienced a larger volume of cricket competition than the sub-elite, up to their international debut age (Jones et al., 2019). We propose that the super-elite's prolonged senior competition experience is partly indicative of the long-term effect of highly dynamic and challenging representative practice offered by higher contextual interference and variable practice, extending the specificity of practice principle, and promoting implicit learning (Henry, 1968; Lawrence et al., 2014; Masters et al., 2008; Pinder et al., 2011). Furthermore, we propose that the super-elite benefited further from the greater exposure to this elite-level competition earlier within their professional careers, given how elite-level competition is likely inherently more representative of international (super-elite) performance than both practice conditions alone, coupled with the lower standard of competition played by the elite during this period.

Transition

From their highest level of amateur club cricket played by age 16, super-elite batsmen transitioned faster than the elite to professional First XI County Cricket; this reflects that they were younger when they made their First XI County Cricket debut, and therefore playing at a higher level of competition from a younger age. The quicker transition rate between competition levels demonstrated by the super-elite mirrors previous research, suggesting that high-potential performers maximize their development from an earlier age, show earlier improvements, and could "make their move" sooner as a result (McCardle et al., 2017). The super-elite's quicker transition does not necessarily denote a "smooth" or linear trajectory into the professional game. Rather, the super-elite's larger volume of challenging practice at age 16, quicker transition to senior competition representation, and extended competition volume thereafter all cumulatively indicate that they were better equipped to deal with the heightened demands of each stage of the pathway, and reflects the optimization of challenge (Ericsson et al., 1993; Guadagnoli & Lee, 2004).

Elite batsmen experienced longer periods of absence from practice and competition due to injury than the super-elite during the early stages of their senior professional county careers (age 19-22). The present finding suggests that elite's higher injury prevalence during this period led to their unavailability for selection on more occasions, and as such, could have contributed to the lower competition volume experienced at ages 21 and 22. The finding represents a "red flag" to science and medicine teams in cricket, given that the super-elite typically made their international debut soon after this period (M_{age} = 23).

Adaptability

The super-elite's superior adaptability was first observed in the second tier of domestic county cricket (Second XI Cricket), who were younger than the elite when they became the best batsmen in their teams. The super-elite were also younger when they became one of the best batsmen in their First XI County team, were more likely to become the best batsman (outright), and were younger when they became the best batsman. These findings offer partial support to two bodies of cricket research, the first demonstrating that elite cricketers achieve their first "significant" performance sooner than sub-elite cricketers. This is strongly correlated with international achievements (Barney, 2015;

Jones et al., 2019). Superior adaptability could be an accelerating factor in transitioning across competition levels, given that instances of this appear as successive occurrences in the super-elite's development timeline (see Coach's Corner). The emergence of longer-term measures of adaptability and the absence of short-term youth performance, as precursors of super-elite expertise within the present findings, highlight the overarching influence of early development experiences, in particular preferences, opportunities, habits, training, and practice activities, as the strongest determinants of sporting mastery. Moreover, the findings suggest that optimizing challenge at a psychological and technical level within practice is a catalyst for the development of (super-elite) batting expertise.

Limitations

The critical reader may identify numerous limitations in the present study. First, as with all self-report retrospective research, the risk of error in recall is attached to findings (e.g., Hopwood, 2013). In an attempt to mitigate this, a matched-pair design was employed in the present study (e.g., Hardy et al., 2013; Güllich et al., 2019); that is, participants were of comparable age, educational background, and cricket playing era (see Method). Furthermore, for questions pertaining to the microdetail of practice (section 4 of the interview schedule), we attempted to alleviate the potential for recall inaccuracies by allowing participants to focus on the season time point (i.e., summer/winter) during which they had reported engaging in the largest volume of practice. Consequently, it was inferred that potential recall inaccuracies owing to age would be approximately equal for both groups. Last, while the interpretation of the 18 discriminating features supports existing theory, it is largely speculative because of the descriptive nature of the research design; we have not explicitly manipulated any variables, but rather used advanced machine learning analysis techniques to classify expertise based on the practice biographies and developmental histories.

Implications for Research and Application

The present study is the first known to have applied a framework to attempt to measure the contextual interference and variability of practice effects in a truly elite sample. The superior predictive power offered by combining (higher) contextual interference and practice variability offers a deliberate practice framework for expertise development in sport; that is, through representing domain-specific practice and providing a mechanism for optimizing challenge simultaneously (Ericsson et al., 1993; Guadagnoli & Lee, 2004). This finding offers a mechanism for which to bridge the limited context specificity posed by deliberate practice theory's conceptualization within a music setting, and its application potential for sport. The super-elite's discriminating random and variable practice observed at age 16 occurred seven years prior to their international debut at age 23. This suggests that research intent on exploring the effects of the microstructure of practice in ecologically valid sporting situations may require more long-term acquisition/practice periods than the short-term effects typically measured in laboratory research. Further examination of factors that moderate the contextual interference effect in sportsmen could lead to a better understanding of the relative contribution of the microstructure of practice in the development of expertise. This represents a fruitful avenue of investigation for experimental research. Above all, the present findings demonstrate that the development of expertise is multifaceted and dynamic. It is therefore imperative that future expertise research extends this holistic approach to identifying precursors of expertise through collecting "wide" data-sets across multiple domains. including psychological and physiological (Jones et al., 2019). The present findings also suggest that the original definition of deliberate practice may not be directly applicable to an elite sporting context (for a review see Ericsson & Harwell, 2019). A modification to the definition of deliberate practice, to describe the nature of practice activity undertaken, rather than enjoyment or satisfaction evoked from the activity, could serve to differentiate deliberate practice from deliberate play better within a sporting context. The suggested modification reflects the fact that the super-elite appeared to enjoy aspects of their random and variable practice implied in the overlap in volume reported across both random and variable practice and play at age 16, which could conceivably be partly due to a specific mindset and/or personality disposition.

In addition to the study's discriminating features, there are 640 features, from the 658 theoretically driven features collected, that do not discriminate between the highest levels of expertise, and can, at the most basic level, be regarded as commonalities (for an overview of all features collected, see Table 2). Several of these commonalities likely contain fundamental developmental experiences that may discriminate between elite and sub-elite batsmen; e.g., undertaking a sufficient volume of blocked practice to develop technical proficiency. Furthermore, while no bowling-related features appeared as discriminators between the elite/super-elite, it could be reasonably hypothesized that facing a sufficient volume of varied bowling types and deliveries during practice, representative of competition, could provide a foundation for preceding levels of expertise (i.e., elite). A replication study is required to test this.

The varying degrees of stringency applied to the analyses have implications for the application of the findings to the field. Specifically, 6 features (from the possible 18 that discriminated) were contained in all 3 solutions derived from the omnibus analyses, demonstrating highest consistency. Three additional features were contained in 2 of the 3 solutions, demonstrating moderate consistency. The remaining 9 features were contained in 1 of the 3 solutions, demonstrating lowest consistency (see Table 5, page 162). Consequently, the authors recommend that the cricket national governing body in England *should act* on features contained in all 3 solutions, *should probably act* on features contained in 2 of the solutions, and *give consideration* to features confined to 1 solution. To understand the complexities of the development profiles of both super-elite and elite batsmen better, a research working group was formulated and was overseen by the corresponding author. This group consisted of three senior ECB officials who have key responsibilities within the talent pathway: head of science, medicine, and innovation; player identification lead; and national lead batting coach. At various stages, expert opinion was sought from these officials. leading to the production of a series of implications and recommendations for talent identification and development, based on the findings. These are presented as part of the "Coach's Corner" (page 182).

Table 5. Level of confidence in feature importance; demonstrated by consensus of features contained within each solution (highest to lowest consistency)

Features Contained in Combined Final Classification Model	Consensus Across Solutions
Volume of random batting practice with maximum variation age 16 Age selected for highest level of cricket competition by age 18 Age became the best batsman in their second XI county team Volume of total cricket activity age 21 (Practice + Competition) Age became the best batsman in their first XI county team	Contained in 3/3 solutions
Age made senior list A (Professional) debut Volume of cricket competition age 22 Age became the best batsman in their first XI county team (Outright)	Contained in 2/3 solutions
Number of older siblings Volume of cricket practice activity age 16 Number of shots practiced randomly age 16 Years to transition from club cricket age 16 to first XI county cricket team Volume of cricket play age 16 Volume of cricket practice activity within busiest practice period age 16 Age selected for highest level county cricket age 22 Development time missed through injury between ages 19-22 Age became one of the best batsmen in their first XI county team	Contained in 1/3 solutions

Conclusion

In conclusion, a pattern of 18 developmental features, from a possible 658, discriminated with excellent accuracy (96%) between super-elite and elite batsmen. Follow-up testing provided evidence of the model's external validity. The overarching influence of challenge represents a foundational difference in the development of super-elite batsmen, compared to the elite, in what appears to be a "race to the top." The super-elite's heightened exposure to older sibling rivalry and associated setback, coupled with their higher degree of contextual interference, indicated by their larger volume of random practice with greater variability, likely equipped them to cope with high-level challenges from an earlier age. This is reflected in super-elite's ability to cope under more challenging circumstances, in the short to medium-term, in transitioning across competition quicker, and adapting to these marked demands sooner than the elite. Their superior long-term skill-retention likely enables the super-elite to develop wider shot strategies and adjust shot parameters, in response to situational information, more effectively in pressurized situations, which represents a performance demand of international cricket.

All considered, the findings suggest that optimizing challenge at a psychological and technical level is a catalyst for the development of super-elite expertise.

Endnotes

1. The term "features" is used to describe groups of variables in this paper.
2. The second and third levels of criteria were determined by the ECB's technical director of elite coaching.
3. This finding reflects that the super-elite were playing at a higher level of competition from a younger age.

Acknowledgments

The authors would like to extend their gratitude to Iain Brunnschweiler, whose technical expertise and feedback were fundamental to the conceptual design of the study and subsequent development of the interview schedule; to all pilot interviewees who gave up their time and provided invaluable feedback; and to Cricket Archive for their data assistance. This research was funded by the England & Wales Cricket Board.

Author's Declarations

The authors declare that there are no personal or financial conflicts of interest regarding the research in this article.

The authors declare that they conducted the research reported in this article in accordance with the Ethical Principles of the Journal of Expertise.

The authors declare that they are not able to make the dataset publicly available but are able to provide it upon request.

References

Baker, J., & Cobley, S. (2013). Outliers, talent codes and myths: Play and practice in developing the expert athlete. In D. Farrow, J. Baker, and C. MacMahon. (eds.) *Developing sport expertise: Researchers and coaches put theory into practice* (pp. 13-25). (2nd ed.). London: Routledge.

Balyi, I., & Hamilton, A. (2004) *Long-term athlete development: Trainability in childhood and adolescence - windows of opportunity - optimal trainability*. Victoria: National Coaching Institute British Columbia & Advanced Training and Performance Ltd.

Barney, E. (2015). *Preliminary stages in the validation of a talent identification model in Cricket* (Unpublished doctoral thesis). Bangor University, UK.

Bishop, C. M. 1995. *Neural networks for pattern recognition* (3rd ed.). Oxford: University.

Brady, F. (2008). The contextual interference effect and sport skills. *Perceptual and Motor Skills, 106*, 461-472.

Broadbent, D., Ford, D. P., O'Hara, D. A., Williams, A. M., Williams, J., & Causer, J. (2017). The effect of a sequential structure of practice for the training of perceptual-cognitive skills in tennis. *PLoS One*,.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, 2,* 121-167.

Côté, J., Baker, J., & Abernethy B. (2007) Practice to play in the development of sport expertise. In Eklund, R., & Tenenbaum, G. (eds), *Handbook of Sport Psychology*. Hoboken, NJ: Wiley, pp. 184-202.

Côté, J., Ericsson , K. A. & Law, M. (2005). Tracing the development of athletes using retrospective interview methods: A proposed interview and validation procedure for reported information. *Journal of Applied Sport Psychology, 17*, 1-19.

Coutinho, P., Mesquita, I., & Fonseca, A.M. (2016). Talent development in sport: A critical review of pathways to expert performance. *International Journal of Sports Science & Coaching, 11*, 279-293.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification,* John Wiley & Sons.

Ericsson, K. A. & Harwell, K. (2019). Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future research. Frontiers in Psychology, 10, 2396.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 3,* 363-406. *Exercise, 11*, 212-222.

Farrow, D., & Buszard, T. (2017). Exploring the applicability of the contextual interference effect in sports practice. *Progress in Brain Research, 234*, 69-83.

Ford, P. R., Low, J., McRobert, A. P., & Williams, A. M. (2010). Developmental activities that contribute to high or low performance by elite cricket batters when recognising type of deliveries from bowlers' advanced postural cues. *Journal of Sport & Exercise Psychology, 32*, 638-654.

Gagné, F. (2004). Transforming gifts into talents: The DMGT as a developmental theory. *High Ability Studies, 15*, 119-147.

Gentile, A. M. (1972). A Working Model of Skill Acquisition with Application to Teaching. Quest, 17, 3-23.

Gladwell, Malcolm. (2011) Outliers: the story of success New York: Back Bay Book.

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, *36*, 212-24.

Güllich, A., Hardy, L., Kuncheva, L., Woodman, T., Laing, S., Barlow, M., Evans, L., Rees, T., Abernethy, B., Côté , J., Warr, C., & Wraith, L. (2019). Developmental biographies of Olympic super-elite and elite athletes: A multidisciplinary pattern recognition analysis. *Journal of Expertise*, *2*, 23-46.

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines, *Machine Learning, 46*, 389-422.

Hall, M. A. (1999). Feature selection for discrete and numeric class machine learning. University of Waikato, *Working Paper Series* 99/4. Correlation-based feature selection for discrete and numeric class machine learning. *Proceedings of the 17th international conference on machine learning (ICML-2000).* 359-366.

Hall, M., Eibe, F., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *SIGKDD Explorations, 11*, 10-18.

Hand, D. J., & Yu, Y. (2001). Idiot's Bayes - not so stupid after all? *International Statistical Review, 69*, 385-389.

Hardy, L., Laing, S., Barlow, M., Kuncheva, L., Evans, L., Rees, T., Woodman, T., Abernethy, B., Güllich, A., Côté, J., Warr, C., Jackson, A., Wraith, L, & Kavanagh, J. (2013). *Great British medallists: A comparison of the biographies of GB super-elite and elite athletes*. End of project report submitted to UK Sport (345 pages).

Hastie, T., Tibshirani, R., & Friedman, J. (2003). *The elements of statistical learning: Data mining, inference, and prediction.* New York: Springer.

Henriksen, K., Stambulova, N., & Roessler, K. (2010). Holistic approach to athletic talent development environments: a successful sailing milieu. *Psychology of Sport and Exercise, 11*, 212-222.

Henry, F. M. (1968). Specificity vs. generality in learning motor skill. In R. C. Brown & G. S. Kenyon (eds.), *Classical studies on physical activity*. NJ: Prentice-Hall.

Hopwood, M. J. (2013). The Developmental History of Athletes Questionnaire: Towards a comprehensive understanding of the development of sport expertise (Doctoral dissertation, Victoria University, Melbourne, Australia). Retrieved from http://vuir.vu.edu.au/id/eprint/22353.

Hopwood, M. J., Farrow, D., MacMahon, C., & Baker, J. (2015). Sibling dynamics and sport expertise. *Scandinavian Journal of Medicine & Science in Sports, 25*, 724-733.

Jones, B. D., Lawrence, G. P., Hardy, L. (2018). New evidence of the relative age effects in "super-elite" sportsmen: A case for the survival and evolution of the fittest. *Journal of Sports Sciences*, *36*, 697-703.

Jones, B. D., Hardy. L., Lawrence, G. P., Kuncheva, L. I., Du Preez, T., Brandon, R., Such, P., & Bobat, M. (2019). The identification of "game changers" within England cricket's pathway for elite spin bowling: A pattern recognition approach. *Journal of Expertise, 2*, 92- 120.

Khan, M.A., Franks, I.M., Elliott, D., Lawrence, G.P., Chua, R., Bernier, P.M., Hansen, S. & Weeks, D. (2006). Inferring online and offline processing of visual feedback in target-directed movements from kinematic data. Neuroscience and Biobehavioral Reviews, 30, 1106-1121.

Kira, K., & Rendell, L. A. (1992). A practical approach to feature selection. In L.Saitta (ed.), *Machine learning: Proceedings of the thirteenth international conference onmachine learning* (pp. 129-134). Aberdeen: Morgan Kaufman.

Kuncheva, L. I. (2014). *Combining pattern classifiers: Methods and algorithm* (2nd ed.) New Jersey: Wiley.

Kuncheva, L. I., & Rodríguez, J. J. (2018). On feature selection protocols for very-low-sample-size data. *Pattern Recognition*, *81*, 660-673.

Lawrence, G. P., Cassell, V. E., Beattie, S., Woodman, T., Khan, M. A., Hardy, L., & Gottwald, V. M. (2014). Practice with anxiety improves performance, but only when anxious: Evidence for the specificity of practice hypothesis. *Psychological Research*, *78*, 634-650.

Lee, T. D., & Magill, R. A. (1985). Can forgetting facilitate skill acquisition? In D. Goodman, R.B.

Wilberg, & I.M. Franks (eds.), *Differing perspectives in motor learning and control* (pp. 283-299). Amsterdam: North Holland.

Lin, C., Fisher, B. E., Winstein, C. J., Wu, A. D., & Gordon, J. (2008). Contextual interference effect: Elaborative-processing or forgetting-reconstruction? A post hoc analysis of TMS-induced effects on motor learning. *Journal of Motor Behavior, 40*, 578-586.

MacNamara, A., Button, A., & Collins, D. (2010). The role of psychological characteristicsin facilitating the pathway to elite performance part 2: Examining environmental and stage related differences in skills and behaviors. *The Sport Psychologist*, *24*, 74-96.

Masters, R. S. W., Poolton, J. M., Maxwell, J. P., & Raab, M. (2008). Implicit motor learning and complex decision making in time constrained environments. *Journal of Motor Behavior, 40*, 71-79.

McCardle, L., Young, B. W., & Baker, J. (2017). Self-regulated learning and expertise development in sport: Current status, challenges, and future opportunities. *International Review of Sport & Exercise Psychology, 12:1*, 112-138. DOI: 10.1080/1750984X.2017.1381141

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*, 134-140.

Pinder, R. A., Davids, K., Renshaw, I., & Araújo, D. (2011). Representative learning design and functionality of research and practice in sport. *Journal of Sport & Exercise Psychology, 33*, 146-155.

Porter, J. M., & Magill, R. A. (2010). Systematically increasing contextual interference is beneficial for learning sport skills. *Journal of Sports Sciences*, *28*, 1277-1285.

Rees, T., Hardy, L., Güllich, A., Abernethy, B., Côté, J., Woodman, T., & Warr, C. (2016). The great British medallists project: A review of current knowledge on the development of the world's best sporting talent. *Sports Medicine*, *46*, 1041-1058.

Rothwell, M., Stone, J. A., Davids, K., & Wright, C. (2017). Development of expertise in elite and sub-elite British rugby league players: A comparison of practice experiences. *European Journal of Sport Science, 17,* 1252-1260.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory 5*, 179-187.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207-217.

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

Yu, L. & Liu, H. (2003). Feature selection for high-dimensional data: A fast correlation-based filter solution. In T. Fawcett & N. Mishra (eds.), *Proceedings of the twentieth international conference on machine learning (*pp. 856-863). Menlo (CA): AAAI.

Submitted: 25 October 2019
Revision submitted: 20 February 2020
Accepted: 16 April 2020

Appendix

COACH'S CORNER

Evidence-Based Recommendations - What to Look For and What to Do

The following recommendations are the result of a consultation between the corresponding author and England & Wales Cricket Board (ECB) talent pathway officials and are intended as a guide for individuals with roles in player identification and player development across the cricket talent pathway in England and Wales:

Player Identification – *What to Look For*

For the identification and selection of county batters:

1. Ask about family and informal cricket play during childhood years (e.g., backyard cricket, competing with older siblings).
2. Look for batters making fast and continual transitions from 15-18 years.
3. Look for batters who become one of the "standout" players in their second and first xi county teams within fewer years.

Player Development – *What to Do*

For batter *program* design:

1. Ensure the appropriate volume of practice is available at age group/academy level.
   Guideline = ~7 hours per week annual average
2. Ensure there is sufficient match play opportunity at age group to academy level.
   Guideline = ~2 matches per week during the summer
3. Ensure there is sufficient opportunity of match days per week at academy to early professional career.
   Guideline = ~100 match days across the calendar year

For batter *practice* design:

1. Ensure a significant proportion of "time on task" is fun and competitive, through a combination of matches, scenario practice, and "net challenges."
   Guideline = > ~50% of total cricket practice time is perceived as "play" by young players
2. Deploy a significant proportion of "random and variable" practice types
   Guidelines =
   - *Split practice time appropriately between the 3 practice levels defined below (Drilling, Mixing It Up, and Scoring Based Scenarios)*
   - *As a guide: Highly random and variable (Scoring Based Scenarios) practice to make up ~40% of skills-based practice time by age 16.*
   - *Ensure that the Mixing It Up and Scoring Based Scenarios practice are as variable as appropriate.*
   - *Keep the challenge level for the player in the "7-8 out of 10" sweet spot, by switching between the levels and/or altering the variability.*

What Does Random and Variable Practice Mean for Coaches?

Defined below are three types of batting practice that coaches can deploy with young players. The first type is "blocked and constant." The second and third types are increasingly "random and variable."

1. Drilling a Specific Shot

- Also known as blocked or fixed practice
- Grooving
- Objective is to become technically proficient at executing a specific shot; e.g., pull shots or front foot drives practiced for 30 minutes.
- Normally involves bowling machine or consistent feeds to similar line and length

2. Mixing It Up

- Also known as random practice
- Develop shot selection and execution
- Objective is to develop the decision-making ability to pick line and length *and* execute a technically sound shot; e.g., mixing between front-foot and back-foot shots to the off-side
- Requires either side arm or real bowling deliveries of various line and length

3. Scoring Based Scenarios

- Also known as "random and variable," "game-based," "net," or "open wicket challenges"
- Objective is to challenge the player to develop and execute run scoring based on game information; e.g., take singles and hit boundaries, over the top or on the ground, to specific areas
- Requires either side arm or real bowling deliveries of various line and length and "field settings" or "target scoring areas"

Summary

- The more varied the practice—in terms of scoring shot options—the greater the challenge and suggested long-term benefit.
- *Health warning:* The super-elite's greater volume of highly random and varied practice (scoring based scenario) practice should not detract from the importance of the other practice types. Each has its own purpose and value; it is important to strike a balance.
- The scoring-based scenarios practice is more "representative" and match-specific. It is therefore essential for performance alongside sufficient technical development from Specific Shot practice and through Mixing It Up practices.