Journal of Expertise
2022. Vol. 5(4)
ISSN 2573-2773

A Retrospective Enquiry into the Holistic Development of Elite British Olympic Weightlifters

Dior N. Jr. Anderson [1,2], Gavin P. Lawrence[1,2], and Victoria Gottwald[1]

[1]Institute for the Psychology of Elite Performance, Department of Sport and Exercise Sciences, Bangor University, Wales, UK
[2]Talent Pathway iD Limited, Wales, UK

Correspondence: Victoria M. Gottwald, v.m.gottwald@bangor.ac.uk

Abstract

Both the scientific community and practitioners continue to be captivated by the challenge of predicting sporting talent but have yet to achieve this with any consistency. Failure can be attributed to oversimplistic models which do not account for the true complexities of talent. Talent is multifaceted, sport specific, and dynamic in nature and should thus be investigated accordingly. The present research adopted a retrospective and multidisciplinary approach to investigate the dynamic development of talent in weightlifting. Using retrospective enquiry, the biographical development of 11 elite and 12 sub-elite athletes was profiled. Developmental themes included the following: (1) demographics and family sport participation; (2) sporting history; (3) competitive milestones; and (4) weightlifting-specific microstructure of practice. A total of 1035 theoretically driven variables were collected on each athlete across their talent developmental lifespan. Phase 1 of the study subjected these data to contemporary machine learning analyses to identify the most predictive pattern of features that best differentiated the elite from the sub-elite sample. The final predictive model needed only six of the 1035 variables to classify the elite and sub-elite with 86% accuracy: total weightlifting related practice by age 12, hours per week flexibility and mobility training by age 14, proportion of whole practice for the clean and jerk by age 19, volume of anxiety- specific practice by age 15 and proportion of anxiety-specific practice by age 19, and the highest international representation by age 19. Phase 2 employed odds ratios as a metric to assess the likelihood of an athlete attaining elite status based on particular attributes. Qualitative accounts of athletes' experiences detailed athletes' transition throughout the pathway. The fundamental components of expertise attainment, and its antecedents, are discussed.

Keywords

Talent identification, talent development, expertise, Olympic Weightlifting, machine learning

Introduction

The development of high performance in sport stems from a dynamic interplay of a multitude of features (Fransen & Güllich, 2019). A considerable body of theoretical frameworks and supporting evidence have illuminated the construct of expertise development over the past 30 years. However, capturing the dynamic interplay between these frameworks and features poses logistical problems for the practitioner and policy maker. Moreover, problems exist when trying to determine which features appear to be more influential than others, particularly when consolidating past research that has (1) predominantly studied factors influencing performance development in relative degrees of isolation and (2) has used statistical approaches that are best suited to experimental and epidemiological research.

In a recently published position stand commissioned by UK sport, the quality of existing evidence from a broad range of factors influencing the attainment of elite sporting performance was explored, and recommendations for policy makers and practitioners were outlined (Rees et al., 2016). Importantly, Rees and colleagues invited research to embrace the complexity and multidisciplinary nature of talent development, rather than continuing to research factors that may influence talent development in isolation. This review has since given rise to a recent body of research that has utilized cutting-edge machine learning data analytics as a solution to this problem (Güllich et al., 2019; Jones et al., 2019, 2020). This machine learning approach has allowed for the selection of a critical set of features in the developmental biographies of athletes that best discriminate between two pre-determined athlete groups (e.g., super-elite versus elite). This critical set can then be used to inform the narrative that best describes the attainment of high performance for the population of interest. To analyze the data in the current investigation, we employed similar state of the art pattern recognition analyses. These analyses were originally used in bioinformatics to identify objects according to features that they possess (Duda et al., 2000). The aim of these analyses is to identify, from a potential large number of features, a subset of features that best discriminate objects of one class from another. More recently, this technique has been used to determine a 'best set' of features from historical data that discriminate between super elite and elite athletes (e.g., GBM study Güllich et al., 2019, DSEB study Jones et al., 2020; DSES study Jones et al., 2019). In the current research, features are the variables we have collected data on (see Table 1 for summary), objects are the athletes from which the data has been collected, and the classes are categories of performance status. This technique has been developed specifically for the analysis of 'short and wide' data sets (i.e., data sets that contain more factors than objects). The analytic procedures involve three distinct processes: feature selection, classification, and finally recursive feature elimination.

In the feature selection process, the analysis aims to identify which set of features correlate well with the class (group) but have low correlation with one another. This feature subset is then tested on its ability to accurately classify the groups. Finally, the analysis aims to refine the feature classification subset to identify the simplest combination of features that best explains the question of interest. It is considered best practice to run multiple algorithms at both the feature selection and classification stages. The more consistent the feature classification between algorithms the more confidence we can place in the predictive validity of that feature subset. A classification rate of 90% for a model indicates if we presented the model with athletes' data for the features within the model subset, we would correctly classify 90 out of every 100 athletes. While traditionally used in bioinformatics, this analytical approach is consistent with the non-linear and holistic nature of talent development in sport and provides researchers with a method that considers an unlimited number of interactions between features; something that is not possible when using traditional statistical approaches. This approach, underpinned by a theoretically driven framework spanning multiple disciplines, enables for a much richer mechanism for conceptualizing expertise development (Jones et al., 2020). Additionally, since data science techniques are in the advent of big data, the potential breadth in exploring the dynamic development of expertise is now as wide as it has ever been.

While the current body of research using these techniques has shed light on the relative importance of many different theoretically driven factors that develop talent, the samples predominantly explored differences in athletes who were selected from a range of different Olympic sports (e.g., Güllich et al., 2019). As such, the themes that have emerged from the research may not necessarily be best suited to the specific characteristics of a single sport. This is consistent with Bergkamp et al.'s (2018) commentary on Johnston and colleagues' (2018) systematic review of the talent identification literature, whereby they highlight the notion that not all factors will contribute equally to expertise development across sports. To date, only two studies have explored the multidisciplinary determinants of expertise development using this approach within a single sport (Jones et al., 2019; 2020); both in the sport of cricket. Using a retrospective research design, Jones and colleagues investigated the relative contributions of a set of 93 multidisciplinary attributes on the development of elite performance attainment in cricket spin bowlers (Jones et al., 2019) and the relative contributions of a set of 693 on the development of elite performance attainment in cricket batsmen (Jones et al., 2020). The authors found that a subset of 12 of the 93 attributes classified elite spin bowlers, with 100% accuracy and a subset of 18 of the 693 attributes classified elite batsmen with 95% accuracy. Interestingly, both the final models are significantly different to one another and retain the multidisciplinary nature of expertise development that was specific to the domains in question (i.e., spin bowling and batting).

These findings highlight that what works in one sport when predicting talent does not necessarily work in another. Indeed, in the work of Jones et al., that is true for disciplines (positions) within a sport. This emphasizes the need for future research to adopt machine learning methodologies within the sport of interest rather than adopt the approach of "one size fits all" when producing athlete profiles that contain the most important developmental factors relative to others.

While machine learning approaches enable us to identify what are arguably the "game changers" (Jones et al., 2019) of elite development, one limitation of this approach is that attributes present across *all* athletes regardless of group membership, can become somewhat dismissed. However, it could be argued that these have equal if not more importance for athlete development, especially given that only a small proportion go on to reach truly elite status (see Johnston & Baker, 2020).

The aim of the current research incorporated two phases: the first phase of the research used a machine learning approach to holistically profile elite performance in weightlifting. While utilizing the theoretical framework discussed by Rees et al. (2016), Phase 1 of the study used a retrospective recall paradigm to explore the extent to which multidisciplinary features across the biographical development (i.e., demographics and family sport participation, sporting history, competitive milestones, and weightlifting-specific microstructure of practice) of elite senior weightlifting athletes can be used to accurately differentiate them from that of their sub-elite counterparts. Phase 2 employed odds ratios as a metric to assess the likelihood of an athlete attaining elite status based on having the particular attributes identified in Phase 1 of the study. Qualitative accounts of athletes' experiences provided richer detail surrounding athletes' transition through the pathway.

Phase 1

Methods

Participants

Twenty-three current and past senior weightlifting athletes volunteered to take part in the study (11 females, 12 males, mean age = 24.3 ± 5.2). The invitations for these studies were administered from the national governing body for weightlifting in Wales. All athletes were informed that the research project formed part of a national talent identification program for the 2022 Commonwealth Games, and they were informed about the specific aims of the study. Ethical approval for the study was granted by the institutional Ethics Committee. So that athlete identity could be preserved for any qualitative accounts throughout the study, athletes were randomly assigned a code so that only the primary researcher could link data to an individual athlete.

Procedure

All athletes were invited to participate in a structured interview which detailed their developmental experiences throughout their formative years. The interview lasted approximately 3 hours, during which athletes were asked a specific series of questions that covered four broad developmental themes: (1) demographics and familial sport participation, (2) sport participation history and weightlifting-specific involvement, (3) competitive milestones in weightlifting, and (4) weightlifting-specific practice activities. The interview was structured to cover any potential environmental influences in section 1, while athlete-specific developmental experiences were covered in sections 2 through to 4. So that any relevant changes to an athlete's developmental experience over time could be recorded, questions in sections 2 to 4 were repeated in relation to their occurrence at each of the following three developmental stages: by age 12, age 15, and age 19. The ages were set to approximately match the early, middle, and later years of athlete development (Balyi, 2001; Bloom, 1985) and were further identified by Weightlifting Wales leadership staff as key developmental milestones within the sport. Quantitative responses were recorded on a spreadsheet for further processing. Audio recordings of the interviews were also transcribed verbatim and were retained to provide qualitative support for any findings. A total of 1035 features were collected and encoded for each athlete. These features are listed in Table 1 (see the Appendix).

Data Analysis

Machine Learning. Machine learning was implemented in the current study to provide a set of rules from which group membership could be best classified. Machine learning normally follows a two-part process: (1) feature selection, then (2) classification (cf. Güllich et al., 2019; Jones et al., 2019; 2020). Feature selection is the process from which the relative importance of the features in the dataset is determined based on predictive validity of the dataset for classifying group membership. A critical subset of the features, usually of a predetermined size, is then established based on the ordering of each feature's relative importance (the highest *n* ranked features are ultimately selected as the model of *n* size). The second step, termed classification, then utilizes classification algorithms to assign each participant with an expected group membership based on their respective scores on each of the features determined through the feature selection process.

This process normally enables the dimensions in the data to become significantly reduced and subsequently allows for the relationships between the groups to be described best using a critical set of features. The process also allows for the features that do not contain predictive validity to be removed from consideration, allowing for a more efficient and less biased approach to describing the patterns in the data. However, in the context of expertise development, in which many complex dynamics occur, this process can lead to the unselected features being somewhat overlooked, and thus may be discarded as unimportant. This could result in a reductionist description of expertise development which may tell only a part of the story. Additionally, commonalities among the groups, which could serve to describe necessary prerequisites for embarking at the beginning of the pathway to elite performance, could also be ignored.

The current study endeavored to use feature selection in such a way that does not over-reduce the multidisciplinary nature of the dataset. Specifically, the study aimed to establish a set of parameters for each feature that can be estimated from the data to establish a rule for each attribute based on the estimated parameter. A subset of these new rules can then be determined by feature selection, which are then carried forward to classification. The machine learning approach in this study followed a four-part process described in the following section: group classification, parameter optimization, feature selection, and final classification. Analyses were performed using the rWeka package in R (Hornik et al., 2009), which is an R interface for the WEKA machine learning statistical software package (Witten et al., 2011).

Group Classification. All athletes in the sample were tracked in a dataset containing all available competition data from the British Weightlifting national governing organization's website (www.britishweightlifting.org). This dataset included all youth, junior, senior and open competitions from January 1, 2001, to January 31, 2022. In addition, all international competition data in which any British athlete had competed between 2001 and 2022 (representing either Great Britain or their respective home nation), was downloaded from either the International Weightlifting Federation's (www.iwf.com/competition-results) or European Weightlifting Federation's website (https://www.ewfed.com). In order to establish an elite sample of senior athletes in the data, any athlete was classified as elite if they had recorded three or more competition totals (which is the combined maximum loads for the snatch lift and the clean & jerk lift) that fell into the top 80% of all historic British competitive performances at one or more of the following senior competitions: (1) a British senior championships, (2) a continental senior championships, (3) a world senior championships (4) a commonwealth senior games or championships, or (5) an Olympic senior games. These competitions were selected as the highest level of representation for British Weightlifting athletes in the senior age group. This resulted in the classification of 11 athletes as elite (6 female, 5 males, mean age = 25.2 ± 6.2; mean age started coach-led weightlifting practice = 14.4 ± 2.9; mean age started weightlifting competitions = 14.6 ± 2.7), and 12 athletes as sub-elite (5 female, 7 males, mean age = 23.5 ± 3.9; mean age started coach-led weightlifting practice = 15.8 ± 3.3; mean age started weightlifting competitions = 16.5 ± 3.3).

Parameter Optimization. To establish the parameters that would be fed into feature selection, a vector of parameters centered on the mean for the elite group was initialized for each attribute in the dataset. This vector was a sequence of 100 equally distributed parameters starting from 3 standard deviations below the mean for the elite group and ending at 3 standard deviations above the mean for the same group. Following this, a set of logical attributes was generated that corresponded to each observation being either over or under each parameter in the vector. Moreover, each athlete was assigned a 1 if their value for the given attribute was above the parameter specified in the vector, and 0 if their value was below this parameter. Consequently, a total of 100 logical attributes were generated for each original attribute in the data.

In order to determine which of the newly generated logical attributes contained the most predictive power for classifying the two groups, feature selection was then performed for each set of 100 logical attributes using a combined rank of the following four feature selection algorithms: correlation attribute evaluator (CAE), the relief F attribute evaluator (Kira & Rendell, 1992), the support vector machine attribute evaluator (cf. Guyon et al., 2002), and the correlation-based feature selection subset evaluator (CFS; Hall & Smith, 1998). This process essentially resulted in a ranked sequence of these variables by order of predictive power. The logical variable that was identified as containing the most predictive power was then put forward for odds ratio estimations (see Phase 2). This resulted in a newly generated dataset containing logical variables, or rules, for each original attribute in the dataset. Consequently, the 387 original attributes were converted to logical rules that were based on an optimized parameter of the attribute.

Feature Selection and Classification. Although it is generally advisable to consider as wide an interpretation of the athlete's development as possible, a summary model was produced using a Bayesian pattern recognition analysis to determine the final model of features to be put forward to classification. As previously mentioned, the final features can allow for a streamlined interpretation of the data, and thus for any instances in athlete monitoring procedures which require interpretation of the critical features only (e.g., such as that in computer software applications). To create the model, feature selection was performed on all normalized attributes in the data (such that the minimum and maximum values for each attribute was represented as 0 and 1, respectively).

Results

This process determined a model of six features which were grouped into three distinct levels of importance based on their appearance in the top 20 features of all four (extremely important), any three (very important), or any two (important) of the FS algorithms, respectively. Table 2 (see Appendix) shows the features in this final model. It is important to note that it is not any standalone feature, but the combination of features that discriminates between athlete status with the accompanying level of accuracy. The tables and radar plot visualizations are simply a representation of the combination of features, and both are designed to aid interpretation and to remind the reader of the need to consider all factors collectively. Within the subset, features may well be interacting in a series of complex ways when discriminating between athletes. Since these are likely to be beyond even three-dimensional in nature, it is impossible to visualize them graphically.

For the next step in the analysis, the model's ability to differentiate the performance groups was assessed against four different classification algorithms. For this step, four commonly used classification algorithms were used: (1) the Naïve Bayes (cf. John & Langley, 1995), (2) J48 decision tree (cf. Quinlan, 1993), (3) Support Vector Machine (SMO; cf. Platt, 1999), and (4) K-nearest neighbours (Aha et al., 1991). This classification process was performed iteratively using a leave-one-out cross-validation procedure in order to minimize overfitting the findings to the data and thus preserving the generalizability of the resulting model. The result of this classification process can be seen in Table 3 (see Appendix). As Table 3 shows, the model was able to differentiate 86% of the sample across all four classification algorithms successfully. The average sensitivity parameter suggests that the model was able to successfully identify the sub-elite sample with 100% accuracy, while the specificity parameter of .71 suggest that 71% of the elite sample was correctly classified on average. An average area under the curve (AUC) of 0.81 also indicates that this model contains moderate predictive power (cf. Obuchowski et al., 2004). The final model with normalized group means is shown as a radar plot in Figure 1 (see Appendix). As is shown, clear separation exists between the groups on each attribute in the model.

Phase 2

While the value of Phase 1 of the research was in understanding the non-linear and interactive nature of attributes that contribute to expertise development, Phase 2 was targeted toward a finer-grained understanding of the identified attributes in isolation and the provision of some useful statistics for practitioners. Specifically, we employed odds ratios to assess the likelihood of an athlete attaining elite status based on having each of the attributes (and related attributes) identified in Phase 1 of the study. Odds ratios are typically used in the medical domain but considered to be a valuable metric for practitioners to understand what our findings might mean in practice. For example, this approach enables us to determine that if an athlete completes X amount of a particular type of practice by age Y, then they will be Z times more likely to reach elite status. For those less familiar with this approach, qualitative accounts of athletes' experiences also provided a rich account of athletes' transition through the pathway.

Methods

Information pertaining to participants and procedure was the same as the first phase of the study.

Data Analysis

Odds Ratio Estimation. As each respective logical variable in the new (logical) data was in the form of a binary variable (i.e., 1 if the applied rule were true, 0 if false), a characteristic shared with the elite performance variable (i.e., 1 if the athlete achieved three or more totals above the 80^{th} British percentile, 0 if false), odds ratios could be calculated for each logical variable in the data. Odds ratios represent the odds of an outcome given the exposure to a condition and are mainly used to assess the effectiveness of clinical trials. In the current study, odds ratios served as a useful metric to assess the likelihood of attaining elite status as a result of achieving the condition associated with each attribute. It also enabled the assessment of the contribution of each rule associated with each attribute to the attainment of elite performance. Odds ratios represent the probability of elite status as a result of exposure to the logical condition relative to the probability of sub-elite given the same exposure.

Odds ratios were therefore produced for each logical attribute in the data. Odds ratios were adjusted for small samples using the small method, and *p* values and confidence intervals were calculated using the Fischer's exact method. A logical rule was considered a discriminator for elite if the *p* values for the associated odds ratio was below 0.05. Conversely, for any logical rules that did not appear as discriminators, commonalities were determined on the basis that (1) a high proportion of each group (approximately 60% or more) met the condition, and (2) the logical attribute contained theoretical relevance as a commonality. These commonalities among the sample could thus be identified as a necessary baseline condition to become involved in the sport and/or progress to senior level in weightlifting.

Results

Table 4 (see Appendix) shows odds ratio estimations (see column 4) for each attribute. These estimations were employed as a metric to assess the likelihood of an athlete attaining elite status based on having each of the attributes (and related attributes) identified in Phase 1 of the study. For example, an odds ratio of 8.57 for flexibility and mobility training by age 14, suggests that an athlete is 8.57 times more likely to reach elite status if they are completing > 0.56 hours of weekly flexibility and mobility training by this age.

Sporting History and Weightlifting-Related Involvement

Elements of reported weightlifting-related involvement throughout development discriminated elite performance attainment. Evidence for these findings predominantly occurred during what is commonly referred to as the middle years of development (i.e., age 13 - 15; Bloom, 1985; Côté et al., 2003). The first discriminating feature in this section was in relation to the number of hours dedicated to flexibility/mobility training from the age of 14. Five of the 11 elite athletes were completing at least 0.56 hours (approximately equivalent to 30 minutes) of flexibility and mobility training per week, while the sub-elite sample reported completing no flexibility or mobility training a week at age 14. This relationship was also apparent at the age of 16, as four of the 11 high performing athletes were completing at least 2 hours of flexibility and mobility training at this age, while no athlete in the sub-elite sample reported completing these levels of flexibility or mobility training at age 16. One of the elite athletes, C-E, who took part in gymnastics from a young age, attributed her levels of flexibility training to her earlier involvement in the sport of gymnastics:

> Well, I started gymnastics from the age of 5 so I was always doing a lot of flexibility and mobility work. We would also be doing a lot of general conditioning then, but I would definitely say flexibility training was a big part of my gymnastics.

The combined weightlifting specific and related practice activities during the middle developmental years was also an important discriminating feature of weightlifting performance in the current sample. Four of the 11 elite athletes were completing at least 1 hour of combined weightlifting related and specific involvement at age 11, which was not replicated in any of the sub-elite sample. This volume increased to 12.5 hours per week by the age of 16, from which seven of the 11 elite athletes (63%), and just two (16%) of the sub-elite sample were completing this volume. At age 19, four of the 11 elite athletes were completing at least 20 hours of combined flexibility, conditioning, and weightlifting specific technical training each week, which was not replicated in any of the sub-elite sample. When discussing her transition into the later stages of the pathway, J-E said:

> I think the transition to becoming a senior international squad member just after age

> 15 increased my desire and intensity to perform and channel all energy into the sport, but more so the high-performance athletes I was training with motivated me. The training increased in strength building, although we still continued with the other elements of plyometrics, fitness, and conditioning to continue my development.

Competitive Milestones in Weightlifting

The highest level of international representation by age 19 also discriminated performance in the current sample. Seven of the 11 elite sample reported competing in at least a continental youth competition by 19 years of age. Contrary to the ratings of their domestic experiences, four of these elite athletes also reported a psychological challenge of the first exposure to this level of competition of at least 6/10 or more. While documenting their experience of this competition, G-E said:

> [That competition] was not good for me. I got to the venue nearly a kilo overweight. I was pretty much kept in the sauna to lose weight so that was a big learning curve for me as far as comp prep. It was hard not because of the technical level of competition, but because of the physical state I was in leading into the competition. So that was again something I had never experienced. I have never experienced having to [lose that much weight]; it sucked the life out of me, so I would say that was probably the most fatigued I had lifted, and probably one of the most fatiguing [competitions] I have ever lifted at. I had never had to battle that whole physical barrier where everything feels so hard because of the physical state you're in. So, having to concentrate so much on being technically good to make sure you are efficient in your lifts, whereas when you're not in that physical state you can kind of bank on the fact that you're fast or you've got strong legs. [When you're that fatigued] all that is taken away from you and you have to concentrate solely on the technical aspects of the lift.

This would appear to suggest that exposure to the elite international stage offers specific and unique challenges that can only be experienced on the international stage.

Microstructure of Practice

For the clean and jerk, the proportion of whole practice, was a discriminatory feature of performance. Ten out of 11 elite athletes reported proportions of whole practice for the clean and jerk over 21% while just four from 12 of the sub-elite sample reported similar whole practice proportions. When asked about their experiences of whole versus part practice in weightlifting, G-E's response tended to agree with the underlying tenets of the whole versus part literature:

> I'm an advocate of just keep it simple and specific, I think. If you can do ten sessions a week then great but with keeping it like that as well there was lots of other ways, I was getting work done without doing any variations or focusing solely on one movement which would be from the block or from the hang or wherever. I've lifted bigger weights since I've been back doing this type of program than I ever did when I was training full time and doing all the different variations and stuff. So, I think that the variations are important when you are young to a degree, probably not as much as I've done and you could do a little less than that, but ultimately people only get comfortable doing snatch and clean and jerk actually doing snatch and clean and jerk.

The proportions as well as the volumes of anxiety specific practice by 15 and 19 years discriminated elite performance. Four of the 11 elite athletes reported that at least 35% of their overall practice volume was anxiety specific by age 15, which was not applicable to any of the sub-elite sample. This had transpired to the accumulation of more than 164 hours of anxiety-specific practice by age 15, as five of the 11 elite athletes had acquired this volume. By age 19, nine of the 11 elite sample (82%) reported that at least 9% of their overall practice included elements that induced the same emotional responses as experienced during competition. This was reported in just four of the 12 sub-elite athletes (33%). This resulted in a threshold of 342 hours of anxiety-specific practice by 19 years of age. When discussing how her training can induce the specific emotions experienced in competition, H-E mentions that this mainly occurs as the intensity of her training increases:

> When the weights get heavier like we are in blocks now leading up to a comp then the pressure is probably a lot greater. It is self-inflicted but I would probably say it's like a 7 out of 10. I can talk myself out of it, but it's still there.

Discussion

This multidimensional study set out to investigate the extent with which elite performance in weightlifting can be explained holistically. With the aid of machine learning and advanced data handling techniques, this study demonstrated that the attainment of elite performance in weightlifting can be described holistically using a series of empirically derived logical statements. A critical subset of these features was shown to differentiate elite performers from their sub-elite counterparts well beyond the level of chance (i.e., with 86% accuracy). The current findings will be discussed in the context of the framework for expertise development adopted throughout the study.

The attainment of elite performance in weightlifting is indeed characterized by a non-linear and dynamic interplay of features. At the forefront of this dynamic interplay, was the differential exposure to the specific practice activities throughout the athlete's development. Specifically, early weightlifting-specific and related exposure tended to be characterized by participation in flexibility and mobility training, as well as weightlifting-specific technical practice from as early as 12 years of age. Later in development, practice activities tended to be characterized by an increased prevalence of whole practice, particularly for the clean and jerk, as well as practice conditions that incorporated anxiety-related training. While it is clear that the current findings do indeed support the notion that the volume of time dedicated to developing one's own performance through deliberate practice is a fundamental and necessary component of the attainment of elite performance (Baker & Young, 2014; Ericsson et al., 1993), its prevalence in the current sample appears to occur as a result of some important antecedents.

In addition, the differences in athletes' performance attainment could be accounted for by differential exposure to the specific practice activities throughout an athlete's development. Explicitly, early weightlifting-specific and related exposure tended to be characterized by participation in flexibility and mobility training, as well as weightlifting-specific technical practice from as early as 12 years of age. These findings would appear to suggest that an early onset of flexibility training encourages the adaptations deemed most appropriate for elite weightlifting performance. As well as involvement in flexibility-related sports, such as gymnastics, this flexibility training was likely performed alongside strength-based activities, which could have promoted both flexibility and strength-based neuromuscular adaptations in elite athletes at early stages in their development. These findings support the notion of engagement in high volumes of deliberate practice being an important component of elite performance attainment (Baker & Young, 2014; Ericsson et al., 1993), as early exposure to these forms of training ultimately resulted in higher cumulative volumes of practice later in the elite athletes' development and may have meant that a reduced amount of sport-specific practice was needed at adulthood. This is in line with Güllich et al. (2022) who present evidence that adult world-class athletes accumulate less main-sport practice in relation to national-class athletes.

In addition to the antecedents and components of practice activities, the specific developmental experiences encountered were also important features in the development of high performance in weightlifting. The specific implications of these developmental experiences are that they should be challenging enough to the athlete to meet their specific skill level (Guadagnoli & Lee, 2004). As was typically observed, elite athletes tended to be introduced to the highest level of domestic competition very early in their development with little prevalence of technical or psychological challenge. It was not until they were introduced to the international stage, did they meet higher demands of challenge. This suggests that, in order to adapt to the level of challenge encountered, it may be prudent to introduce athletes to the international stage as early as possible. This finding is consistent with the notion of the rocky road paradigm (Collins & MacNamara, 2012), which proposes that a high degree of challenge in the developmental experiences of the athlete, dispersed with periods of adjustment, should foster the appropriate psychological adaptations to stress, resulting in a more robust psychological framework for dealing with adversity (Dienstbier, 1989).

Evidence for the proportions of whole practice as a discriminatory feature of performance was present in the current study. Findings revealed that elite athletes tended to engage in higher volumes of both whole practice than their sub-elite counterparts. This finding would conform to the notion of higher volumes of deliberate practice. Of particular interest, was the finding that elite athletes practiced higher proportions of whole practice for the clean and jerk, particularly at later stages of their development. Specifically, these findings suggest that at least one fifth of practice for the clean and jerk should be practiced as a whole movement at later stages in the athlete's development. In line with the premise of whole practice, this would be in order to promote the motor systems to self-organize throughout the whole movement (Kelso, 1995). This would leave the remainder of the practice for the clean and jerk to be dedicated to practicing as constituent parts, potentially relieving pressure on the body and associated energy systems.

While the benefits of part practice should not be discarded, there are two potential reasons why clean and jerk whole practice should be more beneficial for performance development. The first is that the clean and jerk in itself is a movement comprising two discrete movements, in which the transition between these two distinct movements (as well as the phases within each movement) must be fully integrated in order to promote self-organization of movement. Although beneficial for simplifying the learning process (as well as promoting skeletal muscle adaptions), breaking these movements down into chunks creates further distinction between these movements, and as such increasing the proportion of time spent in part practice would compromise the necessary practice time for the integration of these movements during whole practice (Cohen & Sekuler, 2010). The second reason stems from non-linear pedagogical approaches; specifically representational practice design, whereby the practice prescribed should reflect the demands typically experienced in competition, in this case whole practice (Renshaw et al., 2010).

It is apparent from these findings that practice with conditions that are specific to the demands of competition, both in relation to anxiety and context, are important features for the development of high-performance in weightlifting. This finding lends to the notion of the specificity of practice principle (Henry, 1968), which proposed that the best learning experiences stem from those that most closely approximate the target behavior and environmental context. Moreover, practice conditions that closely meet the demands of competition are proposed to encourage the optimization of available sensory information which is likely to be encountered during competition, even to the extent that transfer to different competition conditions disregards this sensory store, which in turn disrupts performance (Elliott, et al., 1995; Khan & Franks, 2000; Khan et al., 1998; Mackrous & Proteau, 2007). Perhaps pertinent to the sport of weightlifting is the notion that practice with anxiety conforms to the principles of specificity of practice and leads to more robust competition performance (Lawrence et al., 2014; Oudejans, 2008; Oudejans & Pijpers, 2009). Most of these athletes reported that these anxious states were encountered toward the end of their competition phases of training when they were starting to train close to the loads they were expecting to open their competition lifts with. This suggests they were adopting the recommendations of Lawrence et al. (2014) regarding the later timing of anxiety induced practice. While it is widely accepted that talent is a complex, dynamic, and multifaceted construct, more traditional approaches to investigate it have explored themes in relative degrees of isolation. This limits the extent to which the holistic element of talent can be understood, particularly with regards to the interactive nature of attributes, for example the relative proportions of whole versus part practice during development stages can influence the level of representation reached as a young adult. The advent of machine learning now makes it possible for researchers to explore important relationships by deploying algorithms that consider the relative importance of a multitude of attributes simultaneously. Machine learning is particularly valuable when investigating "wide" datasets where sample size may be limited but attributes (or variables) are many. While traditionally used in bioinformatics, this approach is well aligned with the non-linear and holistic nature of talent development in sport and enables us to consider an unlimited number of interactions between variables that are not possible when using traditional statistical approaches. The classification rate (in the current study 86%) also provides us with a clear picture of the accuracy and practical value of our model, whereby we are able to classify correctly 86 out of every 100 athletes as being high or low performing based on just the six attributes identified in our model. While the present study employed odds ratios to assess the likelihood of an athlete attaining elite status based on having a particular attribute, future studies may wish to adopt this procedure but instead use combinations of attributes to better tackle the interactive effects identified in Phase 1 of the study.

Limitations

Findings should be taken in context of the limitations that are underpinned by the study's methodology. The retrospective recall of information, particularly in relation to quantitative data, has potential to be influenced by biases that relate to the developmental perspectives of the athletes. Such biases could result in overreporting of any information that participants would inherently see as important to their development, and the inverse for any non-relevant information. Future research should aim to overcome these limitations by investigating the relevance of some of the findings in current youth and junior athletes, who would potentially be undergoing some of the experiences reported by the senior athletes (see Anderson et al., 2022). While the analysis adopted was a strength of the study, the nature of machine learning and its goal to identify features that best distinguish between two classes of objects meant that athletes were classified into a rather binary vision of performance (i.e., elite [high] or sub-elite [lower] performing) where the reality of high performance is arguably less "black and white." Furthermore, the analytical approach meant that discriminating attributes can be truly understood only in the context of the other attributes investigated. We have tried to discuss potential interactions based on our theoretical framework, but these could be interpreted in different ways. While a strength of machine learning is its value when investigating wide data sets, this does come with an important caveat, whereby the model (or pattern of features identified) for optimal classification is arguably specific to the sample adopted in the study. However, it could be argued that this is also the case for many traditional analyses. Similarly, one might expect a perfect model to have a classification accuracy of 100% but this can lead to what is called "overfitting" whereby the model lacks flexibility and thus has reduced value when used on unseen datasets. Finally, while giving athletes the opportunity to talk about some of the factors collected provided richer information and enhanced our understanding of the practical implications, future authors may wish to adopt more robust qualitative analysis protocols to investigate these things further.

Authors' 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.

Acknowledgements

The authors would like to acknowledge the Knowledge Economy Skill Scholarship (KESS) and Weightlifting Wales for funding and supporting the research. They would also like to recognize Professor Lew Hardy for his early contribution to the methodological development.

ORCID iDs

Victoria M. Gottwald
http://orcid.org/0000-0002-0158-9721

Gavin P. Lawrence
https://orcid.org/0000-0001-9232-6718

References

Aha, D. W., Kibler, D., & Albert, M. K. Instance-based learning algorithms. *Machine Learning 6*, 37-66 (1991). https://doi.org/10.1023/A:1022689900470

Anderson, D. A., Gottwald, V. M., & Lawrence, G. P. (2022). The talent pathway: Using machine learning to capture the dynamic development of expertise. *Frontiers in Sports and Active Living: Elite Sports and Performance Enhancement, 4*, 1-17. DOI 10.3389/fspor.2022.986134

Araújo, D., Davids, K., & Hristovski, R. (2006). The ecological dynamics of decision making in sport. *Psychology of Sport and Exercise*, (6), 653-676. https://doi.org/10.1016/j.psychsport.2006.07.002

Baker, J., & Young, B. (2014). 20 years later: Deliberate practice and the development of expertise in sport. *International Review of Sport & Exercise Psychology*, *7*(1), 135-157. https://doi.org/10.1080/1750984X.2014.896024

Baker, J., & Young, B. (2014). 20 years later: Deliberate practice and the development of expertise in sport. *International Review of Sport and Exercise Psychology*, *7*(1), 135-157. https://doi.org/10.1080/1750984X.2014.896024

Balyi, I. (2001). Sport system building and long-term athlete development in British Columbia. *Development*, 1-16. Retrieved from http://iceskatingresources.org/SportsSystemDevelopment.pdf

Baltzell, A., Caraballo, N., Chipman, K., & Hayden, L. (2014). A qualitative study of the mindfulness meditation training for sport: Division I female soccer players' experience. *Journal of Clinical Sport Psychology*, *8*(3), 221-244. https://doi.org/10.1123/jcsp.2014-0030

Behncke, L. (2004). Mental skills training for sports: A brief review. *Athletic Insight: The Online Journal of Sport Psychology, 6*(1), 1-24.

Bergkamp, T. L. G., Niessen, A. S. M., den Hartigh, R. J. R., Frencken, W. G. P., & Meijer, R. R. (2018). Comment on "Talent identification in sport: A systematic review". *Sports Medicine*, *48*(6), 1517-1519. https://doi.org/10.1007/s40279-018-0868-6

Bloom, B. (1985). *Developing talent in young people*. New York: Ballantine.

Bruner, M. W., Macdonald, D. J., Pickett, W., & Côté, J. (2011). Examination of birthplace and birthdate in world junior ice hockey players. *Journal of Sports Sciences*, *29*(12), 1337-1344. https://doi.org/10.1080/02640414.2011.597419

Brustad, R. J. (1993). Who will go out and play? Parental and psychological influences on children's attraction to physical activity. *Pediatric Exercise Science*, *5*(3), 210-223. https://doi.org/10.1123/pes.5.3.210

Brustad, R. J. (1996). Attraction to physical activity in urban schoolchildren: Parental socialization and gender influences. *Research Quarterly for Exercise and Sport*, *67*(3), 316-323. https://doi.org/10.1080/02701367.1996.10607959

Cohen, N. R., & Sekuler, R. (2010). Chunking and compound cueing of movement sequences: Learning, retention, and transfer. *Perceptual and Motor Skills*, *110*(3), 736-750. https://doi.org/10.2466/pms.110.3.736-750

Collins, D. J., Macnamara, A., & McCarthy, N. (2016). Putting the bumps in the rocky road: Optimizing the pathway to excellence. *Frontiers in Psychology*, *7*(SEP). https://doi.org/10.3389/fpsyg.2016.01482

Collins, D., & MacNamara, Á. (2012). The rocky road to the top. *Sports Medicine*, *42*(11), 907-914. https://doi.org/10.1007/bf03262302

Côté, J. (1999). The influence of the family in the development of talent in sport. *The Sport Psychologist*, *13*(4), 395-417. https://doi.org/10.1177/152700250200300301

Côté, J., Baker, J., & Abernethy, B. (2003). From play to practice: A developmental framework for the acquisition of expertise in team sport.

In *Expert performance in sports: Advances in research on sport expertise* (pp. 89-110). https://doi.org/10.4324/9781315670706

Côté, J., Baker, J., & Abernethy, B. (2007). Practice and play in the development of sport expertise. In R. Eklund & G. Tenenbaum (Eds.), *Handbook of sport psychology* (3rd Edition, pp. 184-292). Hoboken, NJ: Wiley.

Côté, J., Macdonald, D. J., Baker, J., & Abernethy, B. (2006). When “where” is more important than “when”: Birthplace and birthdate effects on the achievement of sporting expertise. *Journal of Sports Sciences*, *24*(10), 1065-1073. https://doi.org/10.1080/02640410500432490

Dempsey, J. M., Kimiecik, J. C., & Horn, T. S. (2016). Parental influence on children’s moderate to vigorous physical activity participation: An expectancy-value approach. *Pediatric Exercise Science*, *5*(2), 151-167. https://doi.org/10.1123/pes.5.2.151

Dienstbier, R. A. (1989). Arousal and physiological toughness: Implications for mental and physical health. *Psychological Review, 96*(1), 84-100. https://doi.org/10.1037/0033-295X.96.1.84

Duncan, S. C., Duncan, T. E., Strycker, L. A., & Chaumeton, N. R. (2004). A multilevel analysis of sibling physical activity. *Journal of Sport and Exercise Psychology*, *26*(1), 57-68. https://doi.org/10.1123/jsep.26.1.57

Elliott, D., Pollock, B. J., Lyons, J., & Chua, R. (1995). Intermittent vision and discrete manual aiming. *Perceptual and Motor Skills, 80*(3_suppl), 1203-1213. https://doi.org/10.2466/pms.1995.80.3c.1203

Ericsson, K. A., Krampe, R. T., Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363-406. https://doi.org/10.1037/0033-295X.100.3.363

Ericsson, K. A. (2003). Development of Elite Performance and Deliberate Practice. In *Expert Performance in Sports*. Human Kinetics. https://doi.org/10.5040/9781492596257.ch-003

Fontana, F. E., Furtado, O., Mazzardo, O., & Gallagher, J. D. (2009). Whole and part practice: A meta-analysis. *Perceptual and Motor Skills*, *109*(2), 517-530. https://doi.org/10.2466/PMS.109.2.517-530

Fransen, J., & Güllich, A. (2019). Talent identification and development in game sports. In *Psychology of high performance: Developing human potential into domain-specific talent.* (pp. 59-92). Washington: American Psychological Association. https://doi.org/10.1037/0000120-004

Fredricks, J. A., & Eccles, J. S. (2005). Family socialization, gender, and sport motivation and involvement. *Journal of Sport and Exercise Psychology*, *27*(1), 3-31. https://doi.org/10.1123/jsep.27.1.3

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*(2), 212-224. https://doi.org/10.3200/JMBR.36.2.212-224

Güllich, A. (2017). International medallists’ and non-medallists’ developmental sport activities-a matched-pairs analysis. *Journal of Sports Sciences, 35*(23), 2281-2288. https://doi.org/10.1080/02640414.2016.1265662

Güllich, A. (2018). Sport-specific and non-specific practice of strong and weak responders in junior and senior elite athletics: A matched-pairs analysis. *Journal of Sports Sciences*, *36*(19), 2256-2264. https://doi.org/10.1080/02640414.2018.1449089

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*(1), 23-46.

Güllich, A., Macnamara, B. N., & Hambrick, D. Z. (2022). What makes a champion? Early multidisciplinary practice, not early specialization, predicts world-class performance. *Perspectives on Psychological Science, 17*(1), 6-29. DOI: 10.1177/1745691620974772

Güllich, A., Kovar, P., Zart, S., & Reimann, A. (2017). Sport activities differentiating match-play improvement in elite youth footballers: A

2-year longitudinal study. *Journal of Sports Sciences*, *25*(3), 207-215. https://doi.org/10.1080/02640414.2016.1161206

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, *46*(1), 389-422.

Hall, M. A. & Smith, L. A. (1998). Practical feature subset selection for machine learning. In C. McDonald (Ed.), Computer Science '98 Proceedings of the 21st Australasian Computer Science Conference ACSC'98, Perth, 4-6 February, 1998 (181-191). Berlin: Springer.

Harbourne, R. T., & Stergiou, N. (2009). Movement variability and the use of nonlinear tools: Principles to guide physical therapist practice. *Physical Therapy*, *89*(3), 267-282. https://doi.org/10.2522/ptj.20080130

Hardy, L., Barlow, M., Evans, L., Rees, T., Woodman, T., & Warr, C. (2017). *Great British medalists: Psychosocial biographies of Super-Elite and Elite athletes from Olympic sports*. *Progress in Brain Research* (1st ed., Vol. 232). Elsevier B.V. https://doi.org/10.1016/bs.pbr.2017.03.004

Henry, F. M. (1968). Specificity vs. generality in learning motor skills. *Classical Studies in Physical Activity*.

Hornik, K., Buchta, C., & Zeileis, A. (2009). Open-source machine learning: R meets Weka. *Computational Statistics*, *24*(2), 225-232. https://doi.org/10.1007/s00180-008-0119-7

Helsen, W. F., Van Winckel, J., & Williams, A. M. (2005). The relative age effect in youth soccer across Europe. *Journal of Sports Sciences*, *23*(6), 629-636. https://doi.org/10.1080/02640410400021310

Hodges, N. J., & Franks, I. M. (2002). Modelling coaching practice: The role of instruction and demonstration. *Journal of Sports Sciences*. https://doi.org/10.1080/026404102320675648

Hopwood, M. J., Farrow, D., MacMahon, C., & Baker, J. (2015). Sibling dynamics and sport expertise. *Scandinavian Journal of Medicine and Science in Sports*, *25*(5), 724-733. https://doi.org/10.1111/sms.12387

John, G. H., & Langley, P. (1995). Estimating continuous distributions in Bayesian classifiers. *Proceedings of the eleventh conference on uncertainty in artificial intelligence*. https://doi.org/10.1109/TGRS.2004.834800

Johnston K, & Baker J. (2020). Waste reduction strategies: factors affecting talent wastage and the efficacy of talent selection in sport. *Frontiers in Psychology, 10,* 1-11. https://doi.org/10.3389/fpsyg.2019.02925

Johnston, K., Wattie, N., Schorer, J., & Baker, J. (2018). Talent identification in sport: A systematic review. *Sports Medicine, 48*(1), 97-109. https://doi.org/10.1007/s40279-017-0803-2

Jones, B.D., Hardy, L., Lawrence, G. P., Kuncheva, L., Brandon, R., Thorpe, G. (2020). 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? *Journal of Expertise*, *3*(2), 145-168.

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

Kelso, J. A. S. (1995). Dynamical patterns: The self-organization of brain and behavior. Cambridge, MA: MIT Press.

Khan, M. A., & Franks, I. M. (2000). The effect of practice on component submovements is dependent on the availability of visual feedback. *Journal of Motor Behavior, 32*(3), 227-240. https://doi.org/10.1080/00222890009601374

Khan, M. A., Franks, I. M., & Goodman, D. (1998). The effect of practice on the control of rapid aiming movements: Evidence for an interdependency between programming and feedback processing. *The Quarterly Journal of Experimental Psychology Section A*, *51*(2), 425-444. https://doi.org/10.1080/713755756

Kira, K., & Rendell, L. A. (1992). A practical approach to feature selection. In *Machine learning proceedings 1992* (pp. 249-256). Morgan Kaufmann.

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*(5), 634-650. https://doi.org/10.1007/s00426-013-0521-9

MacDonald, D. J., Cheung, M., Cote, J., & Abernethy, B. (2009). Place but not date of birth influences the development and emergence of athletic talent in American football. *Journal of Applied Sport Psychology*, *21*(1), 80-90. https://doi.org/10.1080/10413200802541868

MacNamara, Á., Button, A., & Collins, D. (2016). The role of psychological characteristics in facilitating the pathway to elite performance part 2: Examining environmental and stage-related differences in skills and behaviors. *The Sport Psychologist*, *24*(1), 74-96. https://doi.org/10.1123/tsp.24.1.74

Magill, R. A. (2007) Motor learning and control concepts and applications. McGraw-Hill, New York.

Mackrous, I., & Proteau, L. (2007). Specificity of practice results from differences in movement planning strategies. *Experimental Brain Research*, *183*(2), 181-193. https://doi.org/10.1007/s00221-007-1031-z

Moesch, K., Elbe, A. M., Hauge, M. L. T., & Wikman, J. M. (2011). Late specialization: the key to success in centimeters, grams, or seconds (cgs) sports. *Scandinavian Journal of Medicine and Science in Sports*, *21*(6), 1-9. https://doi.org/10.1111/j.1600-0838.2010.01280.x

Ong, N. C. H., & Griva, K. (2017). The effect of mental skills training on competitive anxiety in schoolboy rugby players. *International Journal of Sport and Exercise Psychology*, *15*(5), 475-487. https://doi.org/10.1080/1612197X.2016.1153129

Obuchowski, N. A., Lieber, M. L., & Wians, F. H. (2004). ROC curves in clinical chemistry: uses, misuses, and possible solutions. *Clinical Chemistry*, *50*(7), 1118-1125. https://doi.org/10.1373/clinchem.2004.031823

Oudejans R. R. (2008). Reality-based practice under pressure improves handgun shooting performance of police officers. *Ergonomics*, *51*(3), 261-273. https://doi.org/10.1080/00140130701577435

Oudejans, R. R., & Pijpers, J. R. (2009). Training with anxiety has a positive effect on expert perceptual-motor performance under pressure. *Quarterly Journal of Experimental Psychology*, *62*(8), 1631-1647. https://doi.org/10.1080/17470210802557702

Platt, J. (1999). Fast training of support vector machines using sequential minimal pptimization. In *Advances in kernel methods: Support vector learning*.

Quinlan, J. R. (1993). Combining instance-based and model-based learning. In *Machine learning proceedings 1993*. https://doi.org/10.1016/b978-1-55860-307-3.50037-x

Ranganathan, R., & Newell, K. M. (2013). Changing up the routine: Intervention-induced variability in motor learning. *Exercise and Sport Sciences Reviews, 41*(1), 64-70. https://doi.org/10.1097/JES.0b013e318259beb5

Rees, T., Hardy, L., Güllich, A., Abernethy, B., Côté, J., Woodman, T., Montgomery, H., Laing, S., & Warr, C. (2016). The Great British Medalists Project: A review of current knowledge on the development of the World's best sporting talent. *Sports Medicine*, *46*(8), 1041-1058. https://doi.org/10.1007/s40279-016-0476-2

Renshaw, I., Chow, J. Y., Davids, K., & Hammond, J. (2010). A constraints-led perspective to understanding skill acquisition and game play: A basis for integration of motor learning theory and physical education praxis? *Physical Education and Sport Pedagogy*, *15*(2), 117-137. https://doi.org/10.1080/17408980902791586

Rossing, N. N., Nielsen, A. B., Elbe, A. M., & Karbing, D. S. (2016). The role of community in the development of elite handball and football players in Denmark. *European Journal of Sport Science, 16*(2), 237-245. https://doi.org/10.1080/17461391.2015.1009492

Schmidt, R. A., & Wrisberg, C. A. (2008). Motor learning and performance: a situation-based learning approach. Champaign, Human Kinetics.

Soberlak, P., & Côté, J. (2003). The developmental activities of elite ice hockey players. *Journal of Applied Sport Psychology*, *15*(1), 41-49. https://doi.org/10.1080/10413200305401

Stevenson, C. L. (1990). The early careers of international athletes. *Sociology of Sport Journal*, *7*(3), 238-253. https://doi.org/10.1123/ssj.7.3.238

Stuart, M. E. (2003). Sources of subjective task value in sport: An examination of adolescents with high or low value for sport. *Journal of Applied Sport Psychology*, *15*(3), 239-255. https://doi.org/10.1080/10413200305388

Vaeyens, R., Philippaerts, R. M., & Malina, R. M. (2005). The relative age effect in soccer: A match-related perspective. *Journal of Sports Sciences*, *23*(7), 747-756. https://doi.org/10.1080/02640410400022052

Vilar, L., Araújo, D., Davids, K., & Button, C. (2012). The role of ecological dynamics in analysing performance in team sports. *Sports Medicine*, *42*(1), 1-10. https://doi.org/10.2165/11596520-000000000-00000

Welk, G. J., Wood, K., & Morss, G. (2003). Parental influences on physical activity in children: An exploration of potential mechanisms. *Pediatric Exercise Science*, *15*(1), 19-33. https://doi.org/10.1123/pes.15.1.19

Williams, A. M., & Hodges, N. J. (2005). Practice, instruction and skill acquisition in soccer: challenging tradition. *Journal of Sports Sciences*, *23*(6), 637-650. https://doi.org/10.1080/02640410400021328

Witten, I. H., Frank, E., & Hall, M. a. (2011). Data mining: Practical machine learning tools and techniques, third edition. *Annals of Physics*. https://doi.org/10.1002/1521-3773(20010316)40:6

Wulf, G., Höß, M., & Prinz, W. (1998). Instructions for motor learning: Differential effects of internal versus external focus of attention. *Journal of Motor Behavior*, *30*(2), 169-179. https://doi.org/10.1080/00222899809601334

Wulf, G., & Su, J. (2007). An external focus of attention enhances golf shot accuracy in beginners and experts. *Research Quarterly for Exercise and Sport*, *78*(4), 384-389. https://doi.org/10.1080/02701367.2007.10599436

Xiao Lin Yang, Telama, R., & Laakso, L. (1996). Parents' physical activity, socioeconomic status and education as predictors of physical activity and sport among children and youths: A 12-year follow-up study. *International Review for the Sociology of Sport*. https://doi.org/10.1177/101269029603100304

Zachry, T., Wulf, G., Mercer, J., & Bezodis, N. (2005). Increased movement accuracy and reduced EMG activity as the result of adopting an external focus of attention. *Brain Research Bulletin*, *67*(4), 304-309. https://doi.org/10.1016/j.brainresbull.2005.06.035

Received: 23 April 2022
Revision received: 18 October 2022
Accepted: 29 November 2022

Appendix

Table 1. Features used as part of the multidimensional profiling

Developmental theme	Features measured	Supporting Evidence
1. Demographics and Familial Sport Participation		
1.1 Familial Sport Participation: Parental sport involvement	Mother involvement in sport, mother experience in weightlifting, father involvement in sport, father experience in weightlifting, same sex sibling, older same sex sibling, same sex sibling experience in weightlifting	Brustad, 1993, 1996; Dempsey et al., 2016; Fredricks & Eccles, 2005; Stevenson, 1990; Welk et al., 2003; Xiao Lin Yang et al., 1996);
1.2 Familial Sport Participation: Sibling sport participation	Same sex sibling, older same sex sibling, same sex sibling experience in weightlifting	Côté, 1999; Duncan et al., 2004; Hardy et al., 2017; Hopwood et al., 2015; Stuart, 2003
1.3 Homeplace throughout Development	Population of longest residing homeplace between 6-12 years, population density of longest residing homeplace between 6-12 years, population of longest residing homeplace between 13-15 years, town population of longest residing homeplace between 13-15 years, times relocated throughout development	Bruner et al., 2011; Côté et al., 2006; MacDonald et al., 2009; Rossing et al., 2016
1.4 Schooling	Attended sport school between 6-12 years, attended sport school between 13-15 years, school main place for sport participation between 6-12 years, school main place for sport participation between 13-15 years	Güllich et al., 2019
1.5 Relative Age	Month of birth (1 = January; 12 = December), birth quarter (calendar and school; Q1 = Jan-Mar [calendar], Q1 = Sept – Nov [school]), relative age to nearest aged sibling (in days)	Helsen et al., 2005; Jones et al., 2018; Vaeyens et al., 2005
2. Sport History and Weightlifting-Specific Involvement		
2.1 Sport Involvement (between ages 6 – 12, 13 -15, and 16-19 years)	Years involved in each of the following sports: athletics, badminton, basketball, boxing, cricket, CrossFit, dance, football, golf, gymnastics, handball, hockey, horse riding, martial arts, motorsports, mountain biking, rounders, rowing, rugby, swimming, tennis, diving, trampoline; years between 6 and 12 years involved in individual sports, team sports, and cgs sports; total number of sports; years between 13 and 15 years involved in individual sports, team sports, and cgs sports; total number of sports; years between 16 and 19 years involved in individual sports, team sports, and cgs sports; total number of sports	Baker & Young, 2014; Côté et al., 2003, 2007; Güllich, 2017, 2018; Güllich et al., 2017; Moesch et al., 2011
2.2 Weightlifting-Specific and Related Involvement (between ages 6 – 12, 13 -15, and 16-19 years)	Number of competitions per year, exposure to competition (hours/year), time spent in competition (hours/year), flexibility/mobility training (hours/week), number of months involved in weightlifting training (hours/week), weightlifting specific practice (hours/week), strength & conditioning training (hours/week)	Baker & Young, 2014; Ericsson, 1993; Soberlak & Côté, 2003

(Table 1. continued on next page.)

Table 1. Features used as part of the multidimensional profiling

Developmental theme	Features measured	Supporting Evidence
3. Competitive Milestones in Weightlifting		
3.1 Domestic Representation (by ages 12, 15, and 19)	Highest level of domestic representation, age of first appearance at highest domestic representation level, rank of first appearance at highest domestic representation, technical challenge of highest domestic competition, psychological challenge of highest domestic competition	Collins et al., 2016; Collins & MacNamara, 2012; MacNamara et al., 2016
3.2 International Representation (by ages 12, 15, and 19)	Highest level of international representation, age of first appearance at highest international representation level, rank of first appearance at highest international representation, technical challenge of highest international competition, psychological challenge of highest international competition	Collins et al., 2016; Collins & MacNamara, 2012; MacNamara et al., 2016
4. Microstructure of Practice (at ages 12, 15, and 19 years)		
	Deliberate practice vs play	Baker & Young, 2014; Côté et al., 2007; Ericsson, 1993
	Mental skills training	Baltzell et al., 2014; Behncke, 2004; Ong & Griva, 2017
	Vicarious experiences	Bandura, 1986
	Information conveyed to athlete	Hodges & Franks, 2002; Williams & Hodges, 2005.
	Whole/part practice	Magill, 2007; Schmidt & Wrisberg, 2008; Fontana et al., 2009.
	Constant vs varied practice	Harbourne & Stergiou, 2009; Ranganathan & Newell, 2013
	Specificity of practice	Lawrence et al., 2014; Oudejans, 2008; Oudejans & Pijpers, 2009
	Focus of attention	Wulf & Su, 2007; Wulf et al., 1998; Zachry, Wulf et al., 2005
	Prescriptive vs constraints coaching	Araújo et al., 2006; Renshaw et al., 2010; Vilar et al., 2012

Table 2. List of attributes selected for the summary model with their rating of importance and direction of influence on weightlifting performance, and a conceptual understanding of each attribute

Attribute	Importance level	Direction of influence		Conceptual Understanding
		Elite	Sub-Elite	
Weightlifting related involvement				
1. Total combined weightlifting related and specific practice by age 12	Important	+	-	This involves a combination of sport-specific weightlifting practice, strength and conditioning work, and flexibility and mobility training.
2. Flexibility/mobility training at age 14 (hours per week)	Important	+	-	This involves practice dedicated to enhancing athlete flexibility and mobility; e.g., work based on stretching, yoga, and Pilates.
Whole versus part practice				
3. Proportion of whole practice for the clean and jerk by age 19	Very important	+	-	Olympic Weightlifting competition requires athletes to lift maximum weights via the performance of two dynamic, whole-body movements: the "snatch" and the "clean and jerk" (3 lifts of each). Lifts can be broken down into their isolated parts; e.g., the deadlift, the clean, or the jerk, or performed as a whole movement, transitioning through each phase of the lift. Specifically, the clean and jerk involves (dead) lifting the weight from the floor and using the momentum of the body (hips) to clean (or pull) the bar in an upwards trajectory before dropping the body below the bar so that the weight rests on the shoulders. The final phase of the movement (the jerk) is then completed, and the bar is raised above the head in a single fluid motion to lock the arms out.
Specificity of practice				
4. Volume of anxiety specific practice by age 15	Important	+	-	This type of practice involves manipulating the practice environment to incorporate psychological pressure. This can be done via simulating the competition environment in practice (e.g., limiting lifts to 3 of each, competing against a peer, setting a time constraint, or having a crowd). Similarly, to simulate what might be experienced in competition, coaches might want to build consequences into practice sessions to increase cognitive anxiety and physiological arousal (e.g., after a failed lift).
5. Proportion of anxiety specific practice by age 19	Extremely Important	+	-	
Competitive milestones				
6. Highest international level of representation by age 19	Very Important	+	-	This type of practice involves manipulating the practice environment to incorporate psychological pressure. This can be done via simulating the competition environment in practice (e.g., limiting lifts to 3 of each, competing against a peer, setting a time constraint, or having a crowd). Similarly, to simulate what might be experienced in competition, coaches might want to build consequences into practice sessions to increase cognitive anxiety and physiological arousal (e.g., after a failed lift).

Table 3. Summary statistics for all four classification algorithms

Classifier	Accuracy	Sensitivity	Specificity	Area under ROC curve
Naïve Bayes	91.3%	**1.000**	0.818	0.886
Support Vector Machine	87.0%	**1.000**	0.727	0.864
J48 Decision Tree	82.6%	**1.000**	0.636	0.746
K-Nearest Neighbour	82.6%	**1.000**	0.636	0.758
All Classifiers	85.9%	**1.000**	0.705	0.813

Note. Accuracy = Correctly classified observations / total number of observations. Sensitivity = 1 – false positive rate. Specificity = 1 – false negative rate. Area under ROC curve is a measure of model's ability to correctly distinguish the two groups. ROC = Receiver operating characteristic.

Table 4. Logical attributes with estimated odds ratios for all features (highlighted in bold text), and related features, identified in the data mining model (Phase 1)

	Attribute	Sub-Elite	Elite	OR (95% CI)
	Flexibility and mobility training (hours per week) at the following ages:			
	Age 14 more than 0.56 hours	**0/12 (0%)**	**5/11 (45.5%)**	**8.57 (1.01 - 445)**
	Age 16 more than 2.02 hours	0/12 (0%)	4/11 (36.4%)	6 (0.7 - 319.52)
	Total combined flexiblity/mobility, strength and conditioning, and weightlifting specific practice (hours per week):			
	Age 6 (no minimum)	**0/12 (0%)**	**0/11 (0%)**	**NA**
	Age 7 (no minimum)	**0/12 (0%)**	**0/11 (0%)**	**NA**
	Age 11 more than 1.12 hours	**0/12 (0%)**	**4/11 (36.4%)**	**6 (0.7 - 319.52)**
	Age 12 more than 1.17 hours	**0/12 (0%)**	**4/11 (36.4%)**	**6 (0.7 - 319.52)**
	Age 13 more than 2.66 hours	0/12 (0%)	5/11 (45.5%)	8.57 (1.01 - 445)
	Age 14 more than 4.34 hours	2/12 (16.7%)	8/11 (72.7%)	6.67 (1.59 - 65.39)
	Age 15 more than 8.66 hours	1/12 (8.3%)	8/11 (72.7%)	11 (2.26 - 153.3)
	Age 16 more than 12.49 hours	2/12 (16.7%)	7/11 (63.6%)	4.67 (1.15 - 42.68)
	Age 19 more than 20.01 hours	0/12 (0%)	4/11 (36.4%)	6 (0.7 - 319.52)
Competitive Milestones in Weightlifting	**By age 19**			
	International representation			
	Highest international representation level by 19 at least continental youth	**0/12 (0%)**	**7/11 (63.6%)**	**16.8 (1.96 - 887.56)**
	Psychological challenge of highest international competition by 19 over 6/10	0/12 (0%)	4/11 (36.4%)	6 (0.7 - 319.52)
Microstructure of Practice	**Whole/part practice**			
	For the clean and jerk:			
	By age 19			
	Proportion of clean and jerk practice as parts under 78%	4/12 (33.3%)	10/11 (90.9%)	8 (1.69 - 103.46)
	Proportion of clean and jerk practice as whole movement over 21%	**4/12 (33.3%)**	**10/11 (90.9%)**	**8 (1.69 - 103.46)**
	Volume of clean and jerk part practice over 764.7 hours by age 19	0/12 (0%)	6/11 (54.5%)	12 (1.4 - 621.53)
	Volume of clean and jerk part practice over 197.7 hours by age 19	1/12 (8.3%)	8/11 (72.7%)	11 (2.26 - 153.3)
	Specificity of practice			
	Anxiety specificity:			
	By age 15			
	Proportion of overall practice over 35%	0/12 (0%)	4/11 (36.4%)	6 (0.7 - 319.52)
	Volume of anxiety-specificity training over 164 hours	**0/12 (0%)**	**5/11 (45.5%)**	**8.57 (1.01 - 445)**
	By age 19			
	Proportion of overall practice over 9%	**4/12 (33.3%)**	**9/11 (81.8%)**	**4.8 (1.18 - 43.5)**
	Volume of anxiety-specificity training over 342.14	0/12 (0%)	7/11 (63.6%)	16.8 (1.96 - 887.56)

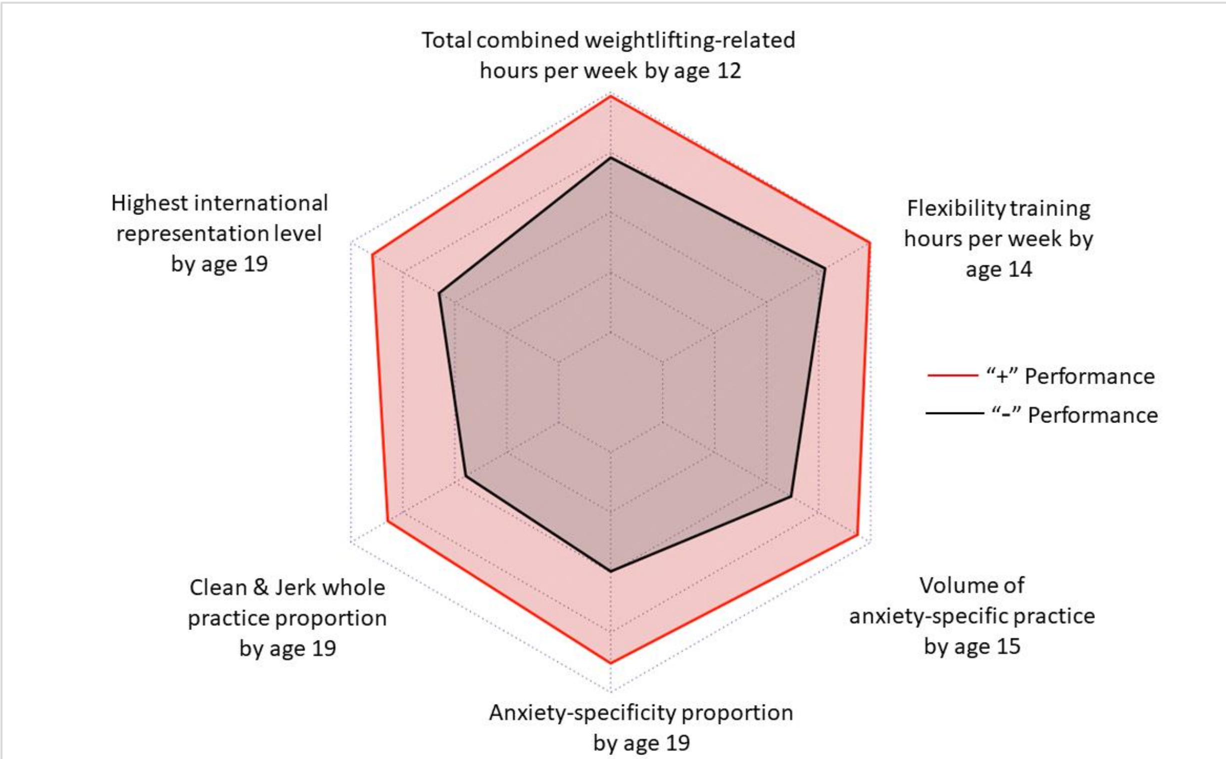

Figure 1. Radar plot depicting the relationships between the performance groups in the summary model