Performing and Observing Complex Skills in Gymnastics: An Investigation of Prototypical Movement Patterns and Perceptual-Cognitive Processes

Von der Sportwissenschaftlichen Fakultät

der Universität Leipzig

genehmigte

DISSERTATON

zur Erlangung des akademischen Grades

Doctor philosophiae (Dr. phil.)

vorgelegt von:

Melanie Mack, M.Sc.

geboren am 01.12.1988 in Ulm

Gutachter:

Prof. Dr. Marco Bortoleto

Prof. Dr. Ivan Çuk

Prof. Dr. Thomas Heinen

Tag der Verleihung: 04. Mai 2021

Abstract

In gymnastics, the performance is evaluated by a subjective rating (Plessner & Haar, 2006), whereby one of the biggest problems is the validity and reliability of judgment (Leskošek et al., 2010; Pajek et al., 2013). The process of observing and evaluating complex motor skills such as those found in gymnastics is located in the cyclical interaction between perception and action (Dosseville et al., 2014; Prinz, 1997), raising the key question, how the quantifiable movement execution is related to the perceived movement quality. Therefore, the thesis aimed at a first step to kinematically analyze selected prototypical gymnastics skills to investigate at a second step the underlying perceptual-cognitive mechanism when observing and evaluating those skills. In the analysis of kinematics, it was focused on the classification and structuring of temporal-spatial continuous kinematic movement patterns and their relationship to the perceived movement quality. In the investigation of the perceptual-cognitive mechanism, it was focused on how skill kinematics are related to the perceived movement quality. Furthermore was the gaze pattern evaluated and examined during the observation. Five consecutive studies were conducted to achieve those objectives.

It was shown that complex gymnastics skills can be kinematically structured into prototypical movement patterns, which differ concerning certain variant and invariant kinematic characteristics. The results of a model-based approach to predict perceived movement quality out of the kinematics of gymnastics skills showed a significant relationship between the predicted score and the true score. Overall, the models worked best for the vault skill, which was the shortest skill with the least spatial variability. Out of all models, the neural network approach showed the best results. Furthermore, it was found, that the similarity of the kinematics of gymnastics skills did not correlate with the similarity of the perceived movement quality. During the evaluation, the gaze pattern was influenced by the manipulation of the amount of non-kinematic information. This was especially the case for the last skill. Expertise seemingly influenced the perceptual-cognitive mechanism partially, but no clear pattern was visible.

The results provided first insights on how gymnastics performances and the resulting judgment scores are related. The complexity of gymnastics skills and additionally, the similarity of the performance level in high elite sport is increasing. This challenge as well as the fast development of technologies leads to the occurrence of computer-based judging systems. Their reliability and validity, but also their challenges and opportunities in artistic gymnastics should be further investigated.

Zusammenfassung

Im Gerätturnen wird die Bewegungsqualität durch ein subjektives Rating bewertet (Plessner & Haar, 2006). Dabei ist Sicherstellung einer hohen Validität und Reliabilität eine grosse Herausforderung (Leskošek et al., 2010; Pajek et al., 2013). Der Prozess der Beobachtung und Bewertung kann in der Interaktion von Wahrnehmung und Handlung verortet werden (Dosseville et al., 2014; Prinz, 1997) und wirft die Frage auf, wie die quantifizierbare Bewegungsausführung mit der wahrgenommenen Bewegungsqualität zusammenhängt. Diese Arbeit zielte in einem ersten Schritt auf die kinematischen Analyse ausgewählter prototypischer turnerischer Elemente ab, um in einem zweiten Schritt die perzeptuell-kognitiven Mechanismen bei der Wahrnehmung und Bewertung zu untersuchen. Bei der kinematischen Analyse lag der Schwerpunkt auf der Klassifizierung und Strukturierung von zeitlich-räumlich kontinuierlichen Bewegungsmustern. Bei der Untersuchung der perzeptuell-kognitiven Mechanismen wurden Zusammenhänge zwischen Kinematik und der wahrgenommenen Bewegungsqualität sowie die Blickmuster während des Beobachtungs- und Bewertungsprozess untersucht. Es wurden fünf aufeinander aufbauende Studien durchgeführt.

Es konnten prototypische Bewegungsmuster identifiziert werden, die sich hinsichtlich bestimmter varianten und invarianten kinematischer Merkmale unterschieden. Die Ergebnisse eines modellbasierten Ansatzes zur Vorhersage der wahrgenommenen Bewegungsqualität durch kinematische Merkmale, zeigten signifikante Zusammenhänge zwischen dem vorhergesagten Wert und dem wahren Wert. Die Modelle lieferten die besten Ergebnisse für das kürzeste Element mit der geringsten räumlich-zeitlichen Variabilität. Über alle Übungen hinweg zeigte der Ansatz der neuronale Netze die besten Ergebnisse. Weiterhin konnte gezeigt werden, dass die kinematische Ähnlichkeit zwischen unterschiedlichen Ausführungen desselben Elements nicht mit der Ähnlichkeit der wahrgenommenen Bewegungsqualität der unterschiedlichen Ausführungen übereinstimmt. Das Blickmuster beim Wahrnehmungs- und Bewertungsprozess konnte durch die Manipulation des Anteils an gezeigter nicht-kinematischer Information beeinflusst werden. Dies war insbesondere beim letzten Element der untersuchten Elementkombination der Fall. Bezüglich des Einflusses von Expertise lassen die Ergebnisse kein klares Muster erkennen.

Die Ergebnisse liefern Erkenntnisse darüber, wie die quantifizierte Bewegungsausführung und die wahrgenommene Bewegungsqualität zusammenhängen. Die Komplexität turnerischer Elemente und die Ähnlichkeit des Leistungsniveaus nimmt zu. Diese Herausforderung sowie die schnelle Entwicklung von Technologien führt zum Aufkommen von computergestützten Beurteilungssystemen. Ihre Güte, aber auch ihre Herausforderungen und Möglichkeiten im Gerätturnen sollten weiter untersucht werden.

List of Publications

In the following, the published articles, included in the PhD thesis, are listed. Parts of the articles may also appear in the Chapters 1-4, 10, 11.

1. Mack, M., Hennig, L., & Heinen, T. (2018). Movement prototypes in the performance of the handspring on vault. *Science of Gymnastics Journal, 10*(2), 245-257. (Chapter 5)
2. Mack, M., Federbusch, S., Ferber, M., & Heinen, T. (in press). Movement prototypes and their relationship in the performance of a gymnastics floor routine. *Journal of Human Sport and Exercise*. (Chapter 6)
3. Mack, M., Schmidt, M., & Heinen, T. (submitted). Mental representation of movement quality of complex skills in gymnastics. (Chapter 7)
4. Mack, M. (2020). Exploring cognitive and perceptual judgment processes in gymnastics by using essential kinematic information. *Advances in Cognitive Psychology, 16*(1), 34-44. (Chapter 8)
5. Mack, M., Bryan, M., Heyer, G., & Heinen, T. (2019). Modeling judges' scores in artistic gymnastics. *The Open Sports Sciences Journal, 12*, 1-9. (Chapter 9)

Contents

Chapter 1

Introduction

Imagine you are a gymnastics judge, and you are selected to score the athletes' performances at the women's gymnastics competition at the 2016 Summer Olympics in Rio. The day of the competition has arrived, and you are on the panel of judges for the women's floor competition. At the moment, Simone Biles, currently known as the best female gymnast in the world, is performing her floor-routine (Figure 1.1). Her second skill combination is a round off - back handspring - double somersault with a half twist, and your full attention is concentrated on her, so that you do not miss any detail of this complicated and fast motion. Less than one second later, she is finishing the skill combination, and you are comparing her movements with the optimal execution model, which is described in the Code of Points (the regular compendium) to find the movement errors and to finally define the performance score. And now, think about the following question: Do you think that you could perceive and remember all the details of the motion? Or could you possibly have missed something, inadvertently, because your visual and cognitive system is not sufficiently precise and quick? I think that everyone, even the best judges in the world, would answer the first question with a resounding "No, I do not think I could perceive and remember all the details!" and the second question with a resounding "Yes, it is possible that I could have missed a movement detail!", and they would agree to the hypotheses, that the fact that Simone Biles is currently the best female gymnast, would influence their decision in such a way, that would they rather judge her performance better than worse; and this because they might have assumed some unsure or unseen movements to have been perfect and error-free.

The judgment of performances in artistic gymnastics has a long tradition. In the early days of gymnastics judging, only one judge evaluated the performances of the athletes. Nowadays, the result of a performance in competition is determined by a panel of judges, who should evaluate a gymnastics performance according to clearly defined rules and

Figure 1.1: Simone Biles, performing her floor-routine at the women's gymnastics competition at the 2016 Summer Olympics in Rio (211UKRjump, 2019, 00:07).

technical requirements. Those guidelines are described in the Code of Points, which nowadays are similar for women and men regarding the criteria, the general evaluation structure, and the judges' panel structure. Each skill is evaluated by six judges, concerning the presentation of the motion, which results in an execution score (E-score), and concerning the skill content, which results in the difficulty score (D-score). Both scores are based upon a 0-10 point scale, whereby the E-score decreases from 10 points in steps of 0.1, and the D-score increases from 0 in steps of 0.1. In a high-level competition, like the Olympic Games or the World Cup, the E-score is determined by six judges and the D-score by two judges. In all lower-level competitions, there are only four judges in the E-score panel. The principal determinant of success is the ability to perform a high-difficulty exercise with a high execution score (Massidda & Calò, 2012). Concerning the E-score, there are four main deductions that could be applied: 0.1 points (a small deduction), 0.3 points (a medium deduction), 0.5 points (a large deduction), and 1.0 point (fall). For instance, if a gymnast performs a skill with bent arms or bent knees, this is penalized with a deduction up to 0.5 points. If the athlete touches the landing mat with the hands during landing, the deduction is 1.0. The sum of the deductions is subtracted from 10.

Judging in artistic gymnastics has a crucial influence on the results. The differences between the performances are often minimal, especially in homogeneous groups, which

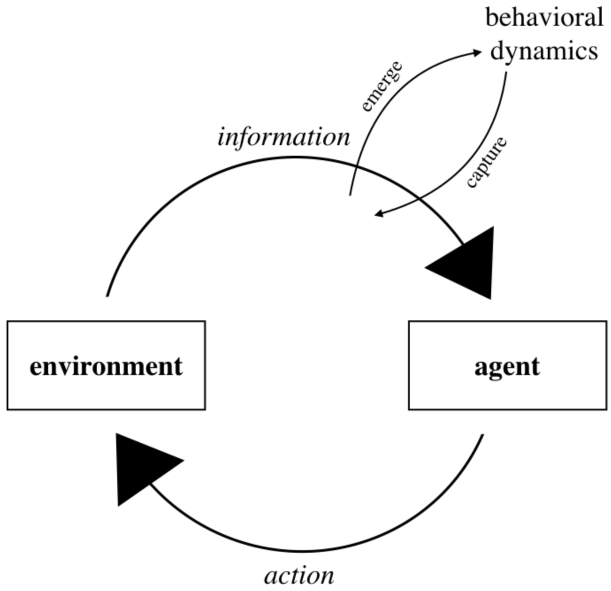

FIGURE 1.2: Schema of the dynamics of perception and action, according to Warren (2006).

could be found for example, in high-level competitions such as European and World Championships or the Olympic Games (GymnasticsResultsCom, 2019). The ranks of an athlete can be influenced by even a small systematic bias of the judges. The variability and the related influencing variables of gymnastics judgments are described and investigated in a large volume of literature, in order to identify and better understand the variables in the evaluation process, where bias could occur. For example, Pajek, Čuk, Pajek, Kovač, and Leskošek (2013) investigated the reliability and validity for female judging at the European championship in Berlin in 2011 - one of the highest level competitions - and they were able to identify inferior reliability of the results for the vault and floor apparatuses. At the University Games, Leskošek, Cuk, Karácsony, Pajek, and Bučar (2010) were able to determine similar results for the vault. It could be assumed that the inferior reliability for the vault is a result of the high velocity of the movement. In the case of the floor apparatus, the judges have to make artistry deductions concerning the creativity of the choreography or the interplay with the music. Because these are all highly subjective categories, they are a possible source of additional bias. Variables, such as the judges' mental capacities, personal experiences, expertise, the point of view, the order of presentation, and the subjectivity of the judging category, are possible sources that could influence the evaluation process. (Ansorge, Scheer, Laub, & Howard, 1978; Heinen, Mandry, Vinken, & Nicolaus, 2013; Heiniger & Mercier, 2018; Pizzera, 2012; Pizzera & Raab, 2012; Plessner & Haar, 2006; Plessner & Schallis, 2005; Ste-Marie, 1999, Ste-Marie, Valiquette, & Taylor, 2001).

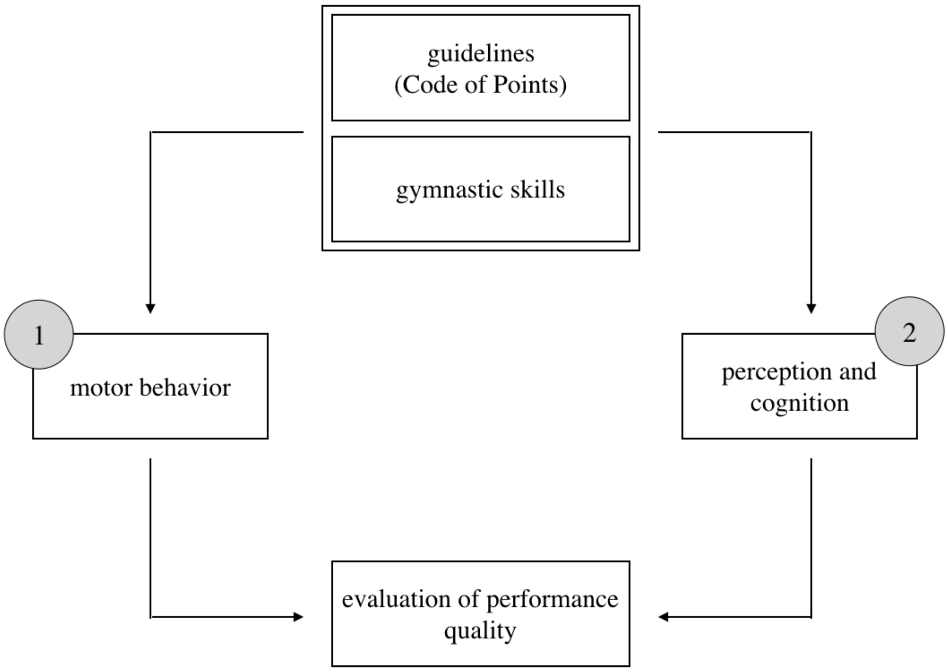

FIGURE 1.3: Relevant theoretical components of the project and the structure of the theoretical background.

However, not only for judges, but also for coaches, gymnasts, and even spectators, the perceived performance quality plays an integral role, albeit differently. While gymnasts use observation to enhance their own ability to move, coaches try to identify errors, thereby deriving useful feedback and training instructions. Judges observe movements concerning defined criteria in the Code of Points, deducing the judgment score from there. Also, other observers, such as spectators, have a feeling of whether a skill is presented in good or bad performance quality. Thereby, one central question arises: how is the quantifiable execution of movement related to the evaluation of performance quality? This relationship is not trivial. It could be shown that very different observers (e.g. judges vs. spectators) come to the same performance score (Čuk, 2015; Ste-Marie, 2003). Besides that, the perception and evaluation of those fast, dynamic, and complex gymnastics skills seem to be error-prone. This might be a problem in the context of a functional and targeting derivation of training instructions and a fair evaluation by judges in competition. The relevance and the actuality of this research become apparent when one considers, that the International Gymnastics Federation is also dealing with this matter by intensifying their efforts to implement a complex judging support system, based on movement data, to support the demand for an objective and potentially unbiased judgment of gymnastics performance (The International Gymnastics Federation & Fujitsu Limited, 2017).

The complex process of observation and evaluation of gymnastics movements can be located within the interaction of perception and action (Dosseville, Laborde, & Garncarcyk, 2014; Prinz, 1997). Thereby, in each task, the agent and its environment are two separate dynamic systems that are informationally and mechanically coupled (Warren, 2006) (Figure 1.2). Concerning the action, this is taken as the quantitative and qualitative analysis of biomechanical variables of the motion. Thereby, the detection of variant and invariant characteristics of the movement structure is of particular relevance. Concerning the perception, this is taken as the quantitative and qualitative analysis of the perceptual-cognitive processes and especially how expertise is mapped therein (Janelle & Hillmann, 2003). Perceptual-cognitive expertise is thereby defined as the ability to identify relevant information in the environment, and to associate it with prior knowledge (Gegenfurtner, Lehtinen, & Säljö, 2011; Mann, Williams, Ward, & Janelle, 2007; Williams & Ericcson, 2005). The aim of this project is, therefore, to kinematically analyze selected prototypical gymnastics skills, in order to investigate, in a second step, the underlying perceptual-cognitive mechanism when perceiving and evaluating those skills. In particular, the relationships between kinematics, gaze pattern, and the estimated performance quality should be focused upon. Based on the results of the projects, relevant recommendations are expected for the practice of gymnastics (and related technical acrobatic sports), which could result in innovative training programs for improvement of the perception and evaluation. This could further enhance the quality of the evaluation of skills and the feedback training, and also contribute to a fair and objective judgement of performances in gymnastics.

Figure 1.3 illustrates the components relevant to the thesis, and their related structures. Because the evaluation of gymnastics skills can be located in the interaction of perception and action, the theoretical background is structured in those two aspects: motor behavior (Figure 1.3, 1) and perception and cognition when evaluating those skills (Figure 1.3, 2). The overall goal thereby was to review the literature concerning models and approaches that do not ignore mechanisms of perception and cognition in the literature on motor behavior, and which do not ignore the mechanism of motor behavior in the literature on perception and cognition. This thesis is therefore structured as follows:

- Chapter 2: Motor Behavior
- Chapter 3: Perception and Cognition
- Chapter 4: Research Questions and Publication Plan
- Chapter 5 - 9: Publications
- Chapter 10: Discussion
- Chapter 11: Conclusion

Chapter 2

Motor Behavior

2.1 Gymnastics Skills as Complex Systems

Artistic gymnastics is a sport that includes highly complex and technically demanding sequences of skills that require maximum effort and a high level of functional ability, such as agility and coordination (Arkaev & Suchilin, 2004). In contrast to most other sports, which consist of a single skill, or just a few skills, in artistic gymnastics, there are six events for men alone (floor, vault, rings, pommel, parallel bars, and horizontal bar) and four events for women (floor, vault, beam, and uneven bars) (Figure 2.1), each of which requires a large number of skills (FIG, 2017). In general, all elements in gymnastics can be divided into static and dynamic skills. The static skills are determined by balance, flexibility, and strength. On a biomechanical level, there is less interest in, and therefore less research into those skills. The dynamic elements, which comprise the majority of the skill repertoire, are determined by coordination and dynamic strength (Brüggemann, 1994). To date, there are different methods of subcategorization, for example, concerning the initial and final positions (Leirich & Rieling, 1969) or concerning the fundamental aims of the movement (Hochmuth & Marhold, 1978).

Across all these categories, gymnastics skills can be seen as complex systems consisting of many components, which interact among themselves, and with the environment as a whole. These interactions change, depending on the constraints embedded within the complex system, and without being previously developed and imposed on the behavior of the system. This movement variability and the resulting coordination dynamics in complex systems tend to form patterned behavior (synergies), which have time-dependent characteristics (Davids et al., 2014). For instance, the movement of the legs while doing a handspring on the vault influences the movement of the trunk, due to interacting joint torques, and the height of the vault or the hardness of the springboard influence the

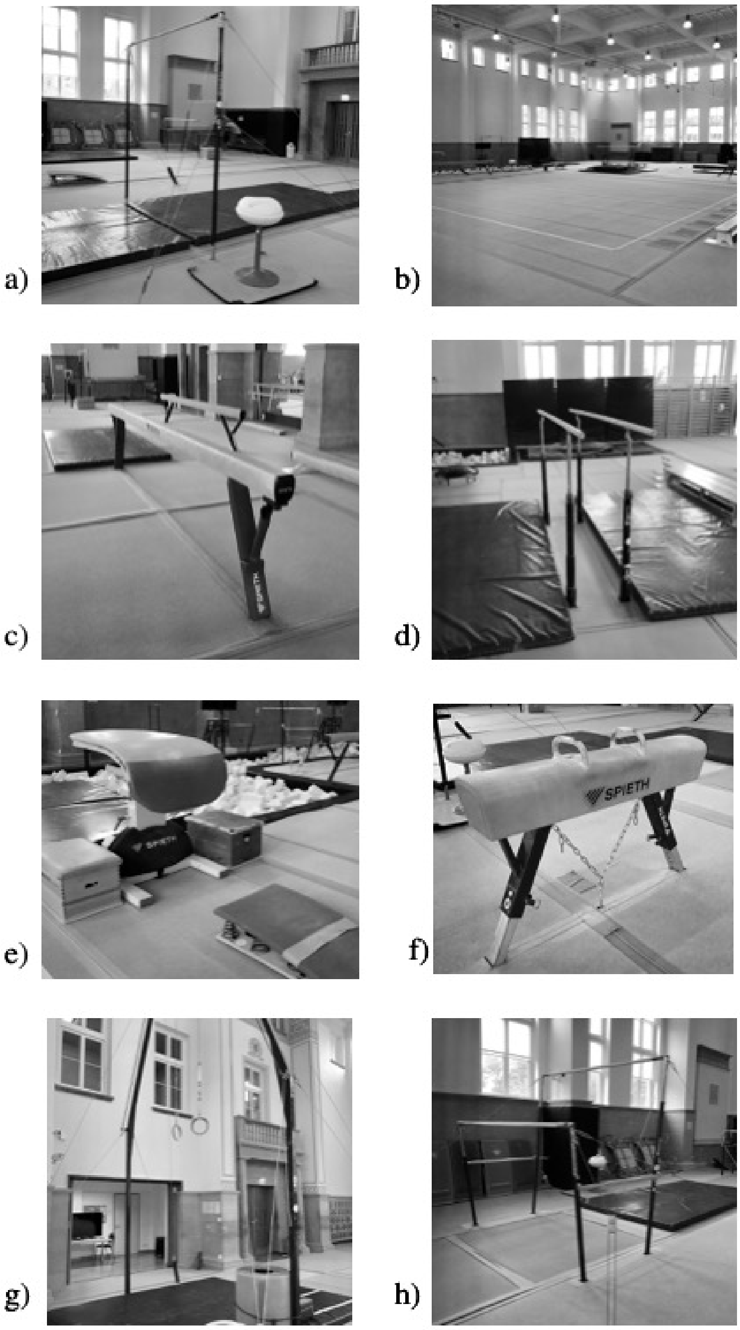

FIGURE 2.1: The six artistic gymnastics events for men: a) horizontal bar, b) floor, d) parallel bars, e) vault, f) pommel, g) rings, and the four events for women: b) floor, c) balance beam, e) vault, h) uneven bars.

handspring as a whole. Also, skills that are performed in a sequence might interact with each other (King & Yeadon, 2004). Thus, for a better understanding of gymnastics performances, the time course of the whole sequence of skills should be taken into account, rather than spatial or temporal-discrete parameters (Troje, 2002).

Due to the presence of the wide range of degrees of freedom in the human motor system, movements can be performed in many different ways (Bernstein, 1967; Latash, Scholz, & Schöner, 2002). There is recent evidence from empirical research that indicates that movement variability is an essential feature of human motor behavior. It affords the necessary flexibility and adaptability to operate proficiently in a variety of performances in fine and gross motor skills (Fitzpatrick, Schmidt, & Lockman 1996; Kelso, 1995; Li, van den Bogert, Caldwell, van Emmerik, & Hamill, 1999), and also for complex skills in gymnastics, comprising whole-body rotations (Hiley, Wangler, & Predescu, 2009; Williams, Irwin, Kerwin, & Newell, 2015). When engaged in a sporting activity, like a basketball shoot, a javelin throw, or a gymnastics skill, for instance, there are different distinguishable movement options that could be described by means of a specific composition of biomechanical parameters.

It is thought that variation in the structure or the function of complex biological systems interacting with different constraints provided by the task, the environment or psychological factors, leads to those different movement options (Davids, Button, & Bennett, 2008; Higgins, 1977). It could, for example, be found, that there are changes in basketball shooting kinematics as a result of distance to the basket (Miller, 2002; Miller & Bartlett, 1996; Robins, Wheat, Irwin, & Bartlett, 2006), gender (Elliott, 1982), ability (Button, MacLeod, Sanders, & Coleman, 2003; Hudson, 1985; Penrose & Blanksby, 1976;) or shooting accuracy (Miller, 1998). Morriss, Bartlett and Navarri (2001) examined the kinematics in javelin throws and could show that experts use different combinations of arm movements to generate pace. Similar results were found for throwing (Schöllhorn & Bauer, 1998), and discus throwing (Bauer & Schöllhorn, 1997). Therefore, it could be assumed, that there is no optimal movement pattern in tasks of that kind. Thereby, it is crucial to consider that in shooting, the only factor relevant to performance is whether or not the ball goes into the basket, no matter how this performance is achieved. By contrast, in gymnastics, the movement pattern itself is the object of the performance evaluation. Nonetheless, even in gymnastics, some performances look more similar than others, which does not implicitly lead to different judgment scores. Biomechanical research has shown that these characteristics are not only descriptive, but also functional. For instance, the Tkatchev, a release and grasp skill on the high and uneven bars, provides an excellent example of the functional role of the particular biomechanical properties of a complex skill. It could be shown that there was a greater angle of release for different versions of the Tkatchev, as the gymnasts extend the hip joint to reach the release angle (Irwin, Manning, & Kerwin, 2011; Kerwin & Irwin, 2010).

In general, it could be said that, within the possible space of movement patterns, there are detectable and characterizable prototypical movement patterns that are similar in their coordination dynamics (Jaitner, Mendoza, & Schöllhorn, 2001; Schöllhorn, Mayer-Kress, Newell, & Michelbrink, 2009), and these can be changed by learning and training influences (Davids, Button, & Bennett, 2008). Those prototypes can be structured, segmented and classified (Endres, Hess & Burgard, 2012; Ren, Patrick, Efros, Hodgins, & Rehg, 2005; Salamah, Zhang, & Brunnett, 2015).

2.2 Biomechanics and Movement Analysis

Before dealing with the particular methods for analyzing the invariant and variant characteristics of the movement pattern, one should first review the literature concerning the available and actual methods used to analyze human movements. The pattern or the sequence of movements can also be labeled with the term "technique" (Beyer, 1992; Carr, 1997). In order to obtain terminological consistency, the term "movement pattern" will be maintained in the following. The literature is strongly linked to the idea of movement as a complex system, which describes the relative position and orientation of body segments and their coordination during the performance of a skill (Lees, 2002). This, in turn, implies, that a particular movement pattern is characterized by visually perceivable variables. (Lees, 2002). These definitions do not indicate how or whether an action could be evaluated as a "good" or "bad" movement pattern or how they relate to performance (Bartlett, 1999; Hay & Reid, 1988; Lees, 2002).

Despite a large number of studies, the biomechanics of gymnastics skills are still poorly understood (Brüggemann, 1994). Especially the question regarding the relationship between visually perceivable kinematics variables and performance remains unclear. It can only be stated that specific time and space-discrete kinematics variables are determined by performance. In the available literature, there is no clear distinction between the general movement pattern (whole-body sequence of movements) and a specific movement pattern (movements of particular body segments) (Lees, 2002). In gymnastics, where the perceived kinematic pattern might be performance-determined, and perception appears to be a complex process where kinematics cues over time and space have been integrated into the evaluation process (Cutting, Proffitt, & Kozlowski, 1978; Johansson, 1973; Runeson & Frykholm, 1983; Troje, 2002; Troje, Westhoff, & Lavrov, 2005), it seems obvious, when analyzing the motion, that one needs to focus on the whole-body sequence of movements. In this thesis, the goal of movement analysis is not to understand the mechanisms that operate while using a particular movement pattern, but rather to describe them in an initial step. In general, one can distinguish between qualitative, quantitative, and predictive approaches for movement analysis.

Qualitative movement analysis is the subjective interpretation of skills in terms of the performance determining aspects for two purposes: 1) to judge performance in aesthetics sports such as artistic gymnastics, and 2) to provide an appropriate intervention to improve performance. For the observation and evaluation of skill performances, several aids (observational models, evaluation templates, etc; for more information see Lees (2002)) had been developed. Until now, there is a lack of agreement in terms of biomechanical principles and and the conceptual framework for qualitative analysis. Further, it requires a high level of knowledge and experience about both, the performed skill and the underpinning biomechanical principles. Therefor, Lees (2002) pointed out that it seems doubtful whether they will advance any further.

In general, quantitative methods use instrumentation to collect data for the quantification of skill performance as a basis for the understanding of a system or a process. This includes motion data, force data, or electromyographic activity. Because as methods for data collection advance and become more widely available this seems not only an objective but also a practical method. The challenge there is to process all the small details being measured in a way that reflects the essential characteristics of technique. Lees (2002) pointed out that with regard to the studies using quantitative approaches, only a few studies are concerned with quantifying different ways of performing a skill, in order to identify the variables that distinguish prototypical skill performances. Most of them aimed to investigate the mean kinematic or kinetic characteristics of a skill (in relation to performance).

With regard to the objective of this thesis (to investigate, in a further step, the perceptual-cognitive processes when evaluating movement) the examination of the variant and invariant characteristics within the whole movement pattern is of particular relevance, since it can be assumed that the invariant characteristics, in particular, are strongly related to the observation and evaluation of actions (Davids, Button, & Bennett, 2008; Takei, 2007). The determination of invariant characteristics of complex athletic movements requires specific multivariate analysis methods, such as cluster analysis, principal component analysis approaches (Daffertshofer, Lamoth, Meijer, & Beek, 2004; Rein, Button, Davids, & Summers, 2010), or Neural Networks (Lees, 2002).

Cluster analysis or principal component approaches, using relative similarities or proximities of variables as a theoretical statistical basis, match the idea of movement as a unique set of interacting biomechanical properties. A commonly used measure for mathematical comparisons is the Euclidean distance, which represents the numerical distance

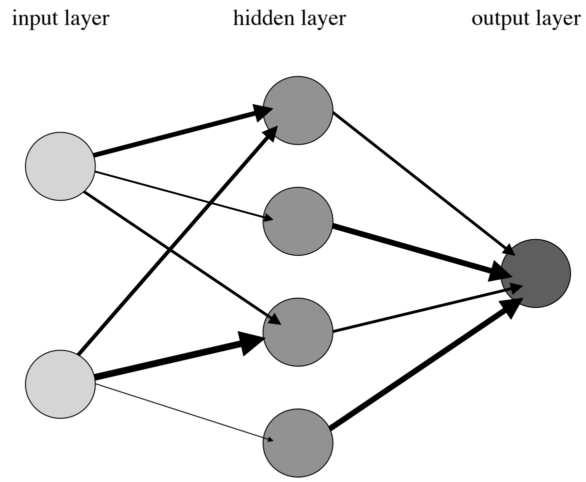

FIGURE 2.2: Functional principle of neural networks.

between two objects (Bauer & Schöllhorn, 1997; Schöllhorn, Chow, Glazier, & Button, 2014). Cluster analysis quantitatively sorts these Euclidean distances, thereby aiming to find groups of objects with a high degree of structural similarity to each other, which can be visualized in a tree diagram (Everitt & Dunn, 2001). Given the natural variation of objects, the different clusters contain a certain degree of variability (Troje, 2002).

The Neural Network approach has only recently been introduced into the analysis of movement patterns. Because of their ability to quantify specific holistic sequences of movement, they thereby offer great potential. Neural Networks are a form of machine learning with a similar model architecture as the human brain. The model can draw relationships between the input variables (i.e. kinematics variables) and the output variables (i.e. performance), based on a non-linear weighting between the input and the output variables (Figure 2.2). There are different applications of neural networks for the analysis of movement patterns. Barton, Lees, Lisboa, and Attfield (2006), for example, used the approach to represent the input variables in a two-dimensional graphical matrix, a so-called Kohonen neural map. The shapes of those visualizations characterize the movement patterns, which in turn means that different shapes signify different movement patterns.

Deterministic models, which can be defined as a graphical or mathematical description of a system or a process, can be used as a basis for the understanding of a system or a process (Chow & Knudson, 2011; Hay & Reid, 1988). In order to identify the presence and the strength of a relationship between a particular variable and a performance outcome, statistical models are used. Thereby, the level of statistical significance is the indicator of a true relationship. For example, research deals with the relationship between the final score and the biomechanical factors of somersault vaults (Farana & Vaverka, 2012; Takei, 1990, 1992, 1998, 2007), Yurchenko vaults (Penitente, Merni, &

Fantozzo, 2009), and Hecht vaults (Takei, Blucker, Nohara, & Yamashita, 2000), or investigates the relationship between the final score and the velocity at the running approach (Naundorf, Brehmer, Knoll, Bronst, & Wagner, 2008; van der Eb et al., 2012; Veličković, Petković, & Petković, 2011). The investigation of changes in these predictive variables is also of interest, in order to evaluate the gymnast's training progress (Bradshaw, Hume, Calton, & Aisbett, 2010). With regard to the objective of this thesis, also for deterministic approaches the focus on the variant and invariant characteristics within the whole movement pattern is of particular interest.

Chapter 3

Perception and Cognition

3.1 Perceptual-Cognitive Processes

Perceptual Expertise in Gymnastics

In sports, expert performance involves both motor and perceptual-cognitive abilities (Williams & Ericsson, 2005). The ability to perceptually determine differences in movement and movement quality in sports, in which the judgment relies on subjective perceptual analysis, has received little attention in research and is poorly understood. It could be assumed, that through the specific and comprehensive qualifications regarding the quality-determining criteria, coaches and judges should gain precise knowledge about the different ways in which skills are executed, in terms of their kinematic features and their matched quality scores. As a result, visual experts, such as coaches or judges, should have a more detailed and precise representation of the optimal execution model, and should, as a consequence, be better at observing movement errors and therefore be able to make a more accurate estimation of the movement quality than motor experts (athletes) or novices (Ste-Marie, Valiquette, & Taylor,2001) can. Studies have shown that it takes sports officials about ten years of extended training to reach the elite level (MacMahon, Helsen, Starkes, & Weston, 2007), which is accompanied by changes in the perceptual-cognitive skills (Ste-Marie, 2000).

In general, perceptual-cognitive expertise is defined as the ability to identify relevant information in the environment, to link it to prior knowledge, to interpret it, and to react appropriately (Gegenfurtner et al., 2011; Mann et al., 2007; Williams & Ericcson, 2005). In an abundance of studies, it could be shown that there are discrepancies in terms of expertise concerning the judgment of performances in gymnastics, but also in the encoding, processing and retrieving of information (Hars & Calmels, 2007; Heinen, Vinken, & Velentzas, 2012; Ste-Marie et al., 2001). Expert judges are better at error detection (Bard, Fleury, Carriere, & Halle, 1980) and perception of body angles (Plessner & Schallies, 2005) compared to novices. The link between perception and action (common coding theory, Prinz, 1990) suggest, that both visual and physical experiences influence the way of perceiving and judging, and should therefore be highlighted and taken into account (Pizzera & Raab, 2012). For instance, it could be shown that the motoric knowledge of a skill leads to a more accurate estimation of joint angles and detection of form errors than purely visual knowledge (Pizzera, 2012). In another study, it could be shown, that gymnasts who were able to perform a skill (but who had no judging experience) judged the skill similarly to gymnastics judges who had judging experience (but no motor experience concerning this particular skill (Heinen et al., 2012). Indeed, both the visual and motor experiences, which judges and coaches have accumulated as athletes and spectators, and in their role as coaches and judges, influence their evaluation performance. Evidence for a generally strong link between perceptual expertise and the kinematics of the observed action originates from studies that demonstrate expert-novice differences in the time-course of information-pick up (Abernethy, Gill, Parks, & Packer, 2001) and anticipatory performances (Abernethy & Zawi, 2007; Ward, Williams, & Bennett, 2002). To explain perceptual expertise in gymnastics, one should first deal with the general processes of human movement perception and discuss how those processes could be measured.

Mental Representation of Movement

It is suggested that actions are organized and stored in the memory as cognitive representations. The perceptual and motor systems thereby share representations for the same action (Prinz, 1997). For instance, when one watches someone performing an action, the motor representation of that action is automatically activated. As a result, it is suggested that those mental structures are linked with the structure of the movement kinematics, and that the mental representation resembles the action in terms of the structural similarity (Glenberg, 1997; Johnsen-Laird, 1989). Those mental representations play a functional role by making situations, objects, and actions cognitively available, although they are physically unavailable, to think and act with them (Vosgerau, 2009). It is proposed that actions are represented in terms of their anticipated perceptual effects (e.g., Prinz, 1997; Hommel, Müsseler, Aschersleben, & Prinz, 2001), and thus, that the mental representation consists of functional features, which are also combined with sensory features that are based on kinematics (Schack, 2004; Schack, 2010; Schack & Ritter, 2009). The functional features refer to action goals that utilize the anticipated effects of the action. The sensory features that are integrated into the action through cognitive chunking refer to the perceptual effects of the movement. Land, Volchenkov, Bläsing, & Schack (2013) stated that this perspective resonates with the idea of Bernstein (1967), suggesting that the interaction of the movement variables do not interact in a fixed manner, but are rather dependent on the movement task and the constraints thereof.

According to Schack's architecture model (Schack, 2004; Schack, 2010; Schack & Ritter, 2009), complex movements are a network of sensorimotor information. Within this network, the related functional subunits, linking the perceptual effects to the goal of the movement, are represented by nodes. One method of analyzing the structure of mental representation is the so-called Structure Dimensional Analysis (SDA). This method was initially developed by Lander and Lange (1996), to analyze the structure of a psychological concept, and it was later adapted to the analysis of movement representation, by Schack (2004; Structure Dimensional Analysis - Motorics, SDA-M). Thereby, the participants have to indicate, whether or not two functional subunits are functionally similar to each other. This is done with all pairs of functional subunits. Based on the participants' decision, the Euclidean distances are calculated between the functional units. Those Euclidean distances can be easily analyzed and visualized by a cluster analysis and the resulting tree-diagram or dendrogram. The SDA-M method has been used in a variety of research, studying expertise concerning cognitive or motor tasks. For example, in tennis, it could be that the structure of the mental representation of expert players was well matched to the biomechanical and functional pattern of the task. On the other hand, the structures of the mental representation of tennis players with less expertise and of novices were less matched to the pattern (Schack & Mechsner, 2006). Similar results could be found for the classical dance movement "pirouette en dehors". The cluster solution of the professional dancers and the advanced amateurs correspond to the functional phases, and the cluster solution of novices differed widely (Bläsing, Tenenbaum, & Schack, 2009). The method was also applied to monitor differences between various development stages (Stöckel, Hughes, & Schack, 2012) or to compare individual cluster solutions to group average cluster solutions (Braun et al., 2007).

In sports such as artistic gymnastics or dancing, not only the anticipated effects of a movement are of importance - the movement quality of a skill as a whole must also be evaluated. From that point of view, not only the representation of a detailed series of action steps are of interest, but also the representation of the relationship of the kinematics variables over the entire time and space during which the movement is performed. To evaluate a gymnastics skill, all biomechanical information has to be summed up and integrated into one evaluation value. The question remains, as to whether movements with a more similar structure of biomechanical parameters were seen as more similar to movements with a less similar structure and whether a similar structure of movement parameters leads to a similar judgment of movement quality. While there is research that deals with searching for the path between biomechanical aspects of movement and the functional elements of movement organization, there is a missing link between biomechanics and the perceived movement quality. Methods, such as the SDA, which deal with the perceived relatedness of two variables, seem to be an approach which could be easily applied to analyze the structure of the perceived quality pattern. After having reviewed the literature concerning how movement is represented in the brain, one should go one level up, to determine how the motion itself is processed.

Motion Perception

"If we see a person we know from far away, while he/she is standing still, it is often difficult to identify that person. Once he/she starts to walk, our visual system receives additional information that increases the likelihood of recognizing that person." (Schöllhorn et al., 2014, pp. 145). This quote describes the nature of perception when observing human beings very vividly. Perception appears to be a complex process with a holistic character, whereby relative information, imparted through the motion itself, is used (Cutting et al., 1978; Runeson & Frykholm, 1983; Troje et al., 2005). Specifically, kinematic cues over time and space, were integrated into the evaluation process (Johansson, 1973; Troje, 2002). An explanation can be found in Johannsson's non-biological motion studies (1975). He found, that relative motion is more important than separate motion. If, for instance, a point of light has a rigid connection to another point of light, those two points of light are perceived as related. It seems that the Gestalt principle "pregnanz", that states that perception moves towards simplicity and wholeness, can also be applied to biological motion perception.

One method used to investigate the influence of kinematics in the evaluation process is the use of point-light displays. Thereby, the major joints of a moving person are represented by dots. Research has shown that not only particular actions (Dittrich, 1993), but also gender (Barclay, Cutting, & Kozlowski, 1978; Troje, 2002) and individual persons can be identified through moving point-light displays (Cutting & Kozlowski, 1977; Stevenage, Nixon, & Vince, 1999; Troje, Westhoff, & Lavrov, 2005). The method was also used in the sporting context, where dynamic information was manipulated or occluded and presented to the participants. For the evaluation, relevant information could be extracted out of the point-light displays (Cañal-Bruland & Williams, 2010; Huys, Smeeton, Hodges, Beek, & Williams, 2008; Williams, Huys, Cañal-Bruland, & Hagemann, 2009). Research has also been done on learning differences when using point-light displays as opposed to videotaped models. Thereby, different results were obtained. Learning through point-light displays was, for example, either less effective (basketball

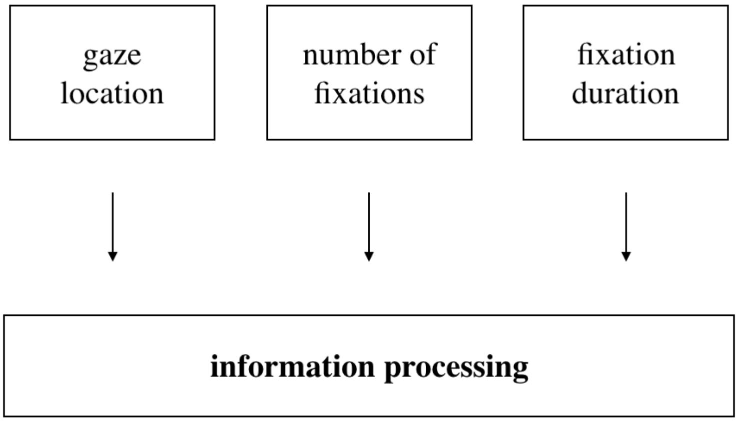

Figure 3.1: Gaze indices, which seems to influence performance.

dribble, Romarck, 1995) or more effective (ballet sequence, Scully & Carnegie, 1998) than learning with original videos.

3.2 Analysis of Perceptual-Cognitive Processes

The investigation of gaze pattern seems to be an appropriate methodology for a better understanding of visual attention and the underlying cognitive processes in the search for information (Vickers, 2007). It can provide essential insights into the information used and the influence of various variables in judgment situations (Bard et al., 1980; Del Campo & Gracia, 2017; Pizzera et al., 2018). Sports performers must be able to identify the areas with the relevant information, direct their attention, extract this information efficiently and effectively, and interpret the meaning thereof (Williams, Davids, & Williams, 1999). Numerous researchers have used eye-tracking systems to record gaze patterns, as athletes attempt to anticipate or judge skill performance, both in the laboratory and in field settings (Mann, Causer, Nakamoto, & Runswick, 2019). Therefore, several indices of gaze could be found, which seem to influence performance by indexing an individual's point of interest and information processing (e.g. the location, the number of fixations, and the duration of fixation, Figure 3.1).

Those indices were used to operationalize certain information-processing assumptions. A longer duration of fixation seems to lead to more detailed information processing (but not necessarily of the fixated visual cues) by allowing the extraction of a higher volume of information from the visual display (Mann et al., 2007). A lower search rate testifies to a reduced information processing load, because of the need of less sensory input, in order to build a coherent mental representation of the display (Abernethy, 1990). The location, the point of time, and the number of fixations could give some indications of selective attention allocation (Gegenfurtner et al., 2011). In a meta-analysis of the underlying mechanism of perceptual skills in sport, Mann et al. (2007) could show that, in general, experts are better at detecting relevant cues by using fewer fixations of longer durations. The results were moderated by variables such as sport type, research paradigm, and presented stimuli. In general, it could be said that skill-based differences in gaze pattern could be found in experts who scan the display more systematically and fixate different areas on the display, than novices do.

In Gymnastics, it could be shown that expertise leads to both a different gaze pattern and different judgment scores. Participants with different sensorimotor gymnastics expertise, focused on different information sources when evaluating gymnastics performances (Bard et al., 1980; Del Campo & Gracia, 2017; Moreno, Reina, Luis, & Sabido, 2002; Pizzera, Möller, & Plessner, 2018; Vickers, 1988). These results suggest that the visual and motor experience influence both the gaze pattern and the information sources on the visual display. Further studies, with a higher amount of participants, should be done to get a clearer picture of the influence of sensorimotor experience on the mechanism when perceiving and evaluating gymnastics skills. Furthermore, in competition, judgment is often not only conducted after a particular skill, but rather, after a sequence of skills (e.g. in a gymnastics floor routine). It would be interesting to investigate the gaze pattern concerning those different skills. Because of the well-known serial-position effect, which is the tendency to recall and remember the first and the last item in a series best, and the intermediate items worst, it could be suggested that there are also differences in the gaze pattern concerning the different items. Because of the importance of kinematics in the evaluation process, and the high amount of information in the visual display, which are not related to those kinematics, it would be interesting to investigate the gaze pattern of gymnasts with different levels of expertise, while evaluating gymnastics movement. Information load could thereby be manipulated, which, for example, has been done in learning studies (Horn, Williams, & Scott, 2002).

Chapter 4

Research Questions and Publication Plan

After having reviewed and discussed the existing literature concerning the aim of this thesis, the following research gaps and research questions could be identified:

1 Most research concerning the kinematics analysis of gymnastics skills only deals with selected temporal or spatial discrete variables. Because of the importance of perceptual-cognitive processes in the evaluation process, it is inevitable that the manifested holistic nature of the perception-process in the analysis be integrated here. Therefore the first aim of the research program was to realize a holistic examination of a gymnastics skill (handspring on vault) as well as a gymnastics skill routine (round off - back handspring - backward layout somersault) with the innovative approach of analyzing temporal-spatial continuous kinematics data. Because of the importance of the variant and invariant characteristics of the movement structure in the evaluation process, these aspects should also be considered. This was done in Chapter 5 and Chapter 6.

2 The questions of whether skills with a similar kinematics pattern were seen as more similar than skills with a less similar kinematics pattern, and whether the similarity is linked to the judgment, remain open. While there is research investigating the path between the functional representation of movements, elements in the brain and the performed movement pattern, there is a missing link between the kinematics pattern and the representation of the movement quality in the brain. Therefore the second aim of the research program was to investigate the relationship between the perceived movement quality of gymnastics skills and its

kinematic pattern, as well as the influence of expertise in this relationship. This was done in Chapter 7.

3 The judgment of gymnastics performances often takes place under conditions of limited time, coupled with restricted resources, and the performance varies widely. There is still a limited understanding of what perceptual information provides the basis for accurate judgment. Therefore, the third aim of the research program was to investigate the perceptual-cognitive processes, and more specifically, the relevance of kinematics in this process. This was done in Chapter 8.

4 Because of the manifold range of influences that could potentially bias the judges' evaluation in artistic gymnastics, a need for alternative approaches, particularly objective computer-based approaches that utilize kinematics data, is claimed. The fourth aim was, therefore, to develop and evaluate a model-based approach to judge gymnastics performances based on quantitative kinematics data. This was done in Chapter 9.

Chapter 5

Publication I

The study described in this chapter has been published following peer review. Full reference details are: **Mack, M., Hennig, L., & Heinen, T. (2018). Movement prototypes in the performance of the handspring on vault.** ***Science of Gymnastics Journal, 10*(2), 227-224.** The author reserves the right that parts are slightly modified.

Abstract

Most research concerning the kinematic analysis of gymnastics skills only deals with selected variables, thereby often ignoring the holistic nature of the analyzed skills. Therefore, the goal of this study was to develop an innovative approach with which to analyze the front handspring on vault. To gain comprehensive insight into the aforementioned motor skill, different skill prototypes should be detected, and their variant and invariant characteristics should be investigated. The digitized video sequences of 60 handspring trials from ten female gymnasts were used for kinematic analysis. Time courses of six joints were analyzed by means of a hierarchical cluster analysis. In addition, the coefficients of variation were calculated. Results revealed that four distinct prototypical movement patterns could be identified for the handspring on vault in female near-expert gymnasts. The movement patterns within each prototype are thereby more similar to each other than the movement patterns between the four prototypes. The four different prototypes can be distinguished by certain variant and invariant characteristics, which become obvious when inspecting the time courses of the hip and shoulder angles, as well as the time course of the coefficient of variation. In light of the training process in gymnastics, the study provides further evidence for strongly considering the gymnasts' individual movement patterns when it comes to the acquisition and optimization of motor skills.

Introduction

Artistic gymnastics involves very complex and technically demanding sequences of elements that require maximum effort and a high level of functional ability, such as agility and coordination (Arkaev & Suchilin, 2004). There are a large number of studies in the field of sport science and gymnastics, that investigate biomechanical aspects of different gymnastics elements (Prassas, Kwon, & Sands, 2006). However, the evaluation of gymnastics performances during training and competition mainly relies on observation by coaches or judges, and is therefore influenced by their respective perception of different kinematic characteristics (Bradshaw & Sparrow, 2001; Farana, Uchytil, Zahradník, & Jandacka, 2015; Farana & Vaverka, 2012). Due to the presence of the high range of degrees of freedom in the human motor system, movements can be performed in many different ways (Bernstein, 1967; Latash et al., 2002). Because humans perceive movement in a holistic way (Davids et al., 2014; Jeraj, Hennig, & Heinen, 2015; Johansson, 1973), the interrelation of movement characteristics with the evaluation can only take place when a movement is taken into account as a whole, and not as a collection of individual parameters. The goal of this study was to identify different prototypes and their variant and invariant characteristics, by means of an innovative approach that allows one to analyze gymnastics skills in a holistic fashion.

Theoretical Background

When engaged in a goal-directed activity, like a handspring on vault, performers develop different coordination states through learning and practice (Davids et al., 2008). Those coordination states are not constantly stable but contain a particular amount of variability, leading to distinguishable movement options that could be described by a specific composition of biomechanical parameters (Latash et al., 2002). Gymnastics skills can be seen as complex systems. Complex systems consist of many components, which interact among themselves and with the environment as a whole. These interactions change, depending on the constraints embedded within the complex system and without being previously developed and imposed on the system's behavior (Davids et al., 2014). The functional role of movement variability in human motor behavior was emphasized in the works of Bernstein (1967) and Higgins (1977), and through nonlinear statistical models in the study of human movement systems (Thompson & Stewart, 2002). It is thought that variation in the structure or function of complex biological systems, interacting with different constraints provided by the task, the environment or psychological factors, leads to movement variability (Davids et al., 2008; Higgins, 1977).

There is recent evidence from empirical research, that movement variability is an essential feature of human motor behavior. It affords the necessary flexibility and adaptability to operate proficiently in a variety of performances, in fine and gross motor skills (Fitzpatrick et al., 1996; Kelso, 1995; Li et al., 1999), and also for complex skills in gymnastics, comprising whole-body rotations (Hiley et al., 2009; Williams et al., 2015). This movement variability and the resulting coordination dynamics in complex systems have a tendency to form patterned behavior (synergies), which have time-dependent characteristics (Davids et al., 2014).

Nowadays, a large amount of kinematic and kinetic data is available to describe human movement. However, sports scientists usually identify, measure, and interpret selected variables, especially on the basis of time-discrete values of selected variables (Federolf, Reid, Gilgien, Haugen, & Smith, 2014; Young & Reinkensmeyer, 2014). Schöllhorn et al. (2014, pp. 145) illustrated the difference between time-discrete and time-continuous movement with the following analogy:

> "If we see a known person far away standing still, it is often difficult to identify that person. Once he/she starts to walk, our visual system receives additional information that increases the likelihood of recognizing that person."

Perception appears to be a complex process with a holistic character that takes into consideration hints and cues that are distributed over the entire time and space, during which the movement is performed, and which is carried both by movement-mediated structural information and by pure dynamics (Troje, 2002). There is further evidence that the perception of biological movement relies on relative movement rather than absolute movement characteristics (Johansson, 1973). Especially in gymnastics, movement is described by coaches, judges, and athletes in terms of specific body postures and movement components (Jeraj et al., 2015). Nevertheless, the challenge is to find an appropriate approach for analyzing the holistic nature of gymnastics skills.

The analysis of quantitative techniques seems unsuitable for establishing the characteristics of the whole skill, but methods, such as cluster analysis or principal component analysis may be able to overcome this limitation (Davids et al., 2014; Lees, 2002). Plausible criteria for the classification of objects seem to be their relative similarity or proximity of movement characteristics. The simplest procedure for classifying objects is to quantify certain characteristics of all objects and to determine the relative distance of these quantities. Joint and body angles seem to be such characteristics, because they can be used to describe gymnastics skills holistically, and other kinematic characteristics can easily be computed from these values (Enoka, 2002).

Hence, qualities can be compared by means of their relative size or vector distance. A commonly used measure for mathematical comparisons is the euclidean distance, which represents the mathematical distance between two objects. Cluster analysis then deals with the quantitative sorting of these euclidean distances (Everitt & Dunn, 2001). If, for example, the euclidean distance between the knee angles of two participants who perform a handspring on vault is smaller than the euclidean distance relative to a third participant, the first two participants would be assigned to one cluster, and the third participant to another cluster. Thus, clustering aims to find groups of objects with a high degree of structural similarity to each other, which can be visualized in a tree diagram. Given the natural variation of objects in relation to their analyzed qualities, the different clusters contain a certain degree of variability (Troje, 2002). In this study, the goal was the identification of prototypical movement patterns of the handspring on vault, by means of a cluster analysis. A prototype is thereby defined by the average angle-time courses of all trials that are assigned to one cluster.

Objectives and Hypothesis

It can be stated that, for a better understanding of complex gymnastics performances, it is not only relevant to analyze isolated parameters, but also to analyze gymnastics skills holistically. Relevant criteria for the classification of objects seem to be their relative similarity or the proximity of kinematic characteristics, like particular joint and body angles. Until now, there has been a lack of gymnastics research that deals with the analysis of gymnastics holistically, based on an explorative approach involving the analysis of time-continuous data. Special interest focused on two topics: (a) to identify prototypes of a gymnastics skill, and (b) to explore the structure of a gymnastics skill in terms of its variant and invariant characteristics.

In a first step, the angle-time courses of separate trials of one specific gymnastics skill (handspring on vault) were mathematically analyzed by means of a cluster analysis. It was hypothesized that some trials are more similar than others. The cluster analysis should reveal patterns of similarity, leading to a particular number of distinguishable clusters (i.e. prototypes). In a second step, the variant and invariant characteristics were investigated qualitatively by analyzing the angle-time courses in relation to the different prototypes and the different movement phases. It was hypothesized that, in specific movement phases, the prototypes differ in terms of their variant and invariant characteristics.

Methods

Participants

Ten female gymnasts participated in this study (age: $M = 11.50$, $SD = 1.43$; body size: $M = 143.00$ cm, $SD = 11.36$ cm). The gymnasts reported an average training duration of 26 hours per week. They were able to perform the experimental tasks required for this study with a high degree of consistency in training and competition (handspring on vault; see Motor task section).

Motor Task

The motor task was a handspring on vault (Čuk & Karácsony, 2004). The vaulting table was arranged according to the competition guidelines of the International Gymnastics Federation for women's artistic gymnastics (FIG, 2017). There was a running track in front of the table, landing mats (0.20 m high) behind the table, and a certified springboard (1.20 m long and 0.60 m wide) in front of the table. The vaulting table was adjusted to a height of 1.25 m.

The handspring on vault can be subdivided into six movement phases: approach run and hurdle, take-off phase, first flight phase, repulsion phase, second flight phase, and landing phase (Brüggemann, 1994). From a standing position at the beginning of the running track, the gymnast performs an accelerated run-up towards the vault apparatus. A hurdle motion at the end of the run-up precedes a reactive leap onto the springboard, which in turn precedes the first flight phase, in which the gymnast supports herself on both hands on the vaulting table. During the support phase, the gymnast pushes herself off the vaulting table, and performs a whole-body rotation around the somersault axis during the subsequent flight phase. The handspring ends with the gymnast landing on both feet in an upright body posture. Gymnasts were asked to perform handsprings on vault, as they would do in a regular competition. In particular, they were asked to perform handsprings with the highest movement quality they were capable of at the time of the study.

Movement Analysis

The performance of the gymnasts was recorded with a digital video camera (240 Hz, 1920 x 1080 pixel), which was placed at a distance of about 15 m from the vaulting table, in order to compensate for lens distortion. The camera videotaped the gymnasts' performance orthogonally to the direction of movement, simulating the judge's perspective. The recorded video sequences were then used for the kinematic analysis. The horizontal and vertical coordinates of 18 points (body landmarks) were digitized for each frame, using the movement analysis software Simi Motion®. Thus, each of the 18 body landmarks was represented by a two-dimensional time series [$x_j(t); y_j(t)$] with $j = 1, 2, 3, ..., j$ (t = time, j = frame number). The 18 body landmarks defined a 17-segment model of the human body (Enoka, 2002). A digital filter that was built into the software was applied for data smoothing. For each trial, the time series of each body landmark was time-normalized and rescaled to the interval [0; 1000]. Kinematic angular data was calculated from the time-normalized position data of the body landmarks for all handspring trials (Jaitner, Mendoza, & Schöllhorn, 2001). The calculated joint angles (knee, hip, shoulder) were specified with regard to the frontal horizontal body axis, thereby reflecting flexion and extension movement (Behnke, 2001).

Procedure

The study was conducted in three phases. In the first phase, the gymnast arrived at the gymnasium. She was informed about the general procedure of the study. In particular, the gymnast was told that she would be taking part in a study on kinematic analysis of the handspring on vault. The study was conducted in compliance with the Helsinki Declaration and the International Principles governing research on humans, as well as in line with the ethical guidelines of the local ethics committee. The gymnast gave her informed consent, and was given a 20-minute warm-up period. After the warm-up, the gymnast was allowed one familiarization trial. In the second phase, the gymnast performed ten handsprings on vault. She was allowed to take breaks as requested and there was no time pressure. In the third phase, and after completing the ten handsprings on vault, the gymnast was debriefed and dismissed into an individual cool-down period.

Data Processing and Analysis

The free statistic software R (R Core Team, 2017) was used for further data processing and analysis. The further data analysis comprised two steps: In the first step, the prototypical movements of the handspring on vault were identified by means of a hierarchical cluster analysis. Therefore, euclidean distances were calculated for each time course of joint angles (see equation 1: x and y denote a corresponding joint angle between a pair of two handsprings, and i denotes a point in the rescaled time interval [0;1000]). The resulting values were summed up to form one euclidean distance value for each pair of two handspring trials. Thereby, a value of zero would have indicated an exactly identical course of two handspring trials, whereas the larger the resulting value, the more dissimilar two trials were.

$$d(x,y) = \sqrt{\sum\nolimits_{i=1}^{n} (x_i - y_i)^2} \tag{5.1}$$

The resulting euclidean distance values were recorded on a distance matrix, indicating the similarity between each pair of two handspring trials. In order to classify all trials by means of their similarity, the euclidean distance matrix was evaluated quantitatively by a hierarchical cluster analysis, using Ward's hierarchical clustering method (Ward, 1963). Ward's method was chosen because it is an agglomerative clustering method that is based on a classical sum-of-squares criterion, producing groups that minimize dispersion within a group at each fusion step (Murtagh & Legendre, 2014). The classification result was represented by a two-dimensional tree diagram, illustrating the fusions or divisions made at each stage of the analysis. The number of clusters was determined by inspecting the scree plot in terms of the elbow criterion (Everitt & Dunn, 2001). In a second step, and in order to characterize each of the prototypes, the time courses of the joint angles were averaged over the corresponding trials in each cluster. In addition, the time courses of the coefficient of variation were calculated for all joint angles of each prototype, indicating the relative extent of variability of a particular prototype along its time course (Stergiou, 2004).

Results

Figure 5.1 presents the result of the hierarchical cluster analysis. Following the inspection of the scree plot of the cluster analysis, four subgroups (i.e., clusters) could be distinguished from each other. Each of the four clusters thus comprised handspring trials that were more similar to trials within a particular cluster, but which were more dissimilar to handspring trials in the other clusters. Therefore, each cluster characterized a particular handspring prototype within the sample of all analyzed handspring trials. Inspecting the individual clusters revealed that prototype #a comprised 17 handspring trials (28.33 %), and prototype #b comprised 15 (25 %) handspring trials. Prototype #c contained 7 handspring trials (11.67 %), and prototype #d contained 21 handspring trials (35 %). A subsequent Chi-square test revealed a statistical trend that indicated that the number of handspring trials was not distributed equally between the four clusters, $\chi^2 = 6.93$, $p = .07$, indicating that some handspring prototypes appear more frequently

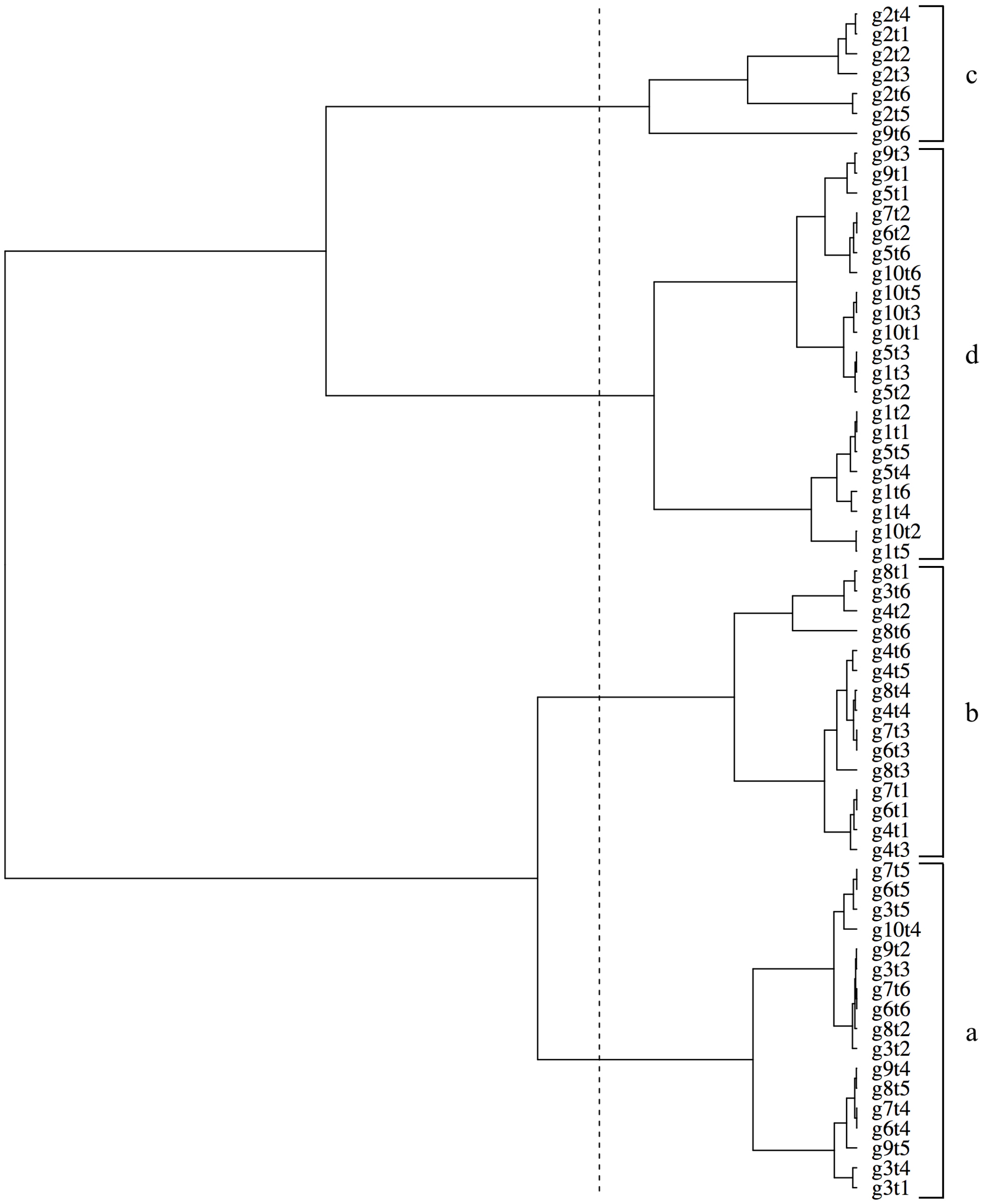

FIGURE 5.1: Tree diagram resulting from a cluster analysis using Ward's clustering algorithm. Horizontal lines indicate the level of the distance at which the respective handspring trials are grouped into one cluster. *Note*: The dashed line represents the euclidean distance below which the clusters are identified. The letters *a* to *d* correspond to the four clusters, containing the different prototypical movement patterns of the handspring on vault, *g1t1* to *g10t6* represent the analyzed handspring trials.

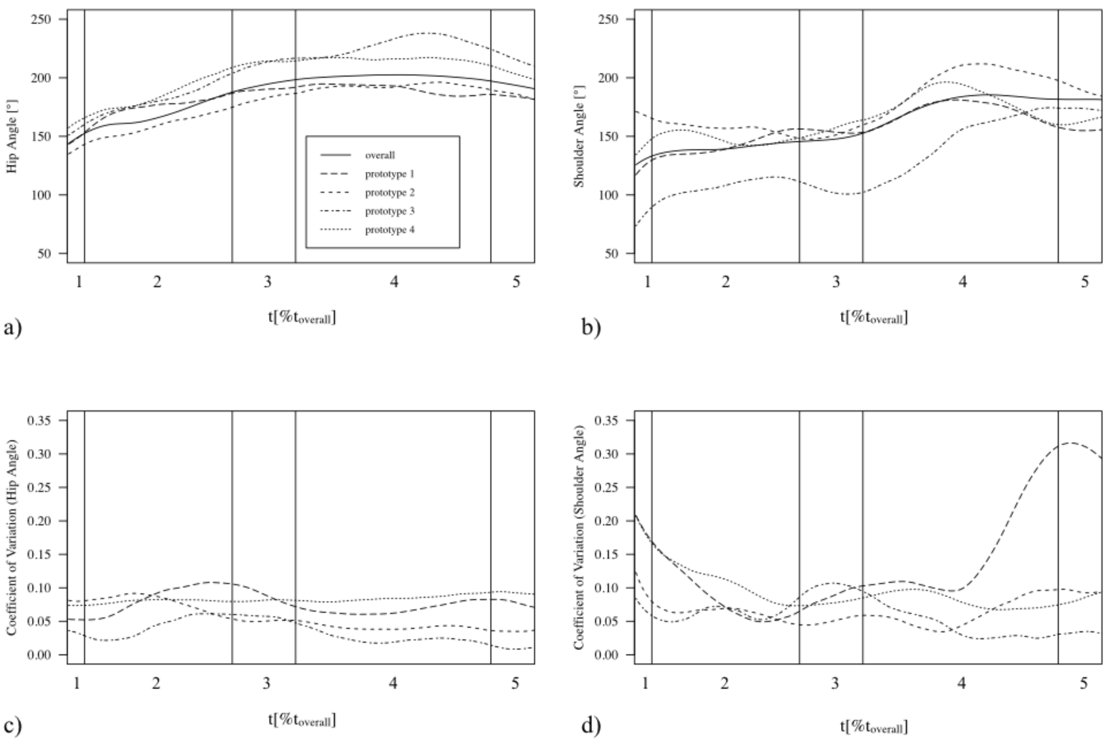

FIGURE 5.2: Illustration of time-normalized angle-time plots for the prototypical courses of the shoulder angle (a) and hip angle (b), as well as time courses of the corresponding coefficients of variation for the different prototypes (c, d). *Note:* 1 = take-off phase from springboard, 2 = first flight phase, 3 = repulsion phase, 4 = second flight phase, 5 = landing phase.

in gymnasts, such as prototype #d, while other prototypes appear less frequently, such as prototype #c.

Prototype 1

In cluster #a, 17 trials were grouped together. A picture sequence of an exemplary trial can be seen in Figure 5.3 a). Exemplary time courses of hip and shoulder joints can be found in Figure 5.2 a) and 5.2 b). A typical handspring trial from cluster #a comprised the following characteristics: 1) slightly inclined trunk with open shoulder angle during touch-down on springboard, 2) inclined trunk, slightly flexed hip joint and open shoulder angle during take-off from the springboard, 3) slightly flexed hip joint, open shoulder angle and trunk orientation close to 45° during touch-down on the vaulting table, 4) slightly overarched back, and stretched hip and shoulder joints during take-off from the vaulting table, and 5) straight back with slightly flexed hip and knee joints and open shoulder angle during touch-down on the landing mat. There was a rather small coefficient of variation for the hip joint and the knee angle over the time course (0 - 0.1). For the shoulder angle, the coefficient of variation was about 0.2 at the take-off phase which decreased to 0.1 during the first flight phase. Exemplary time courses of the

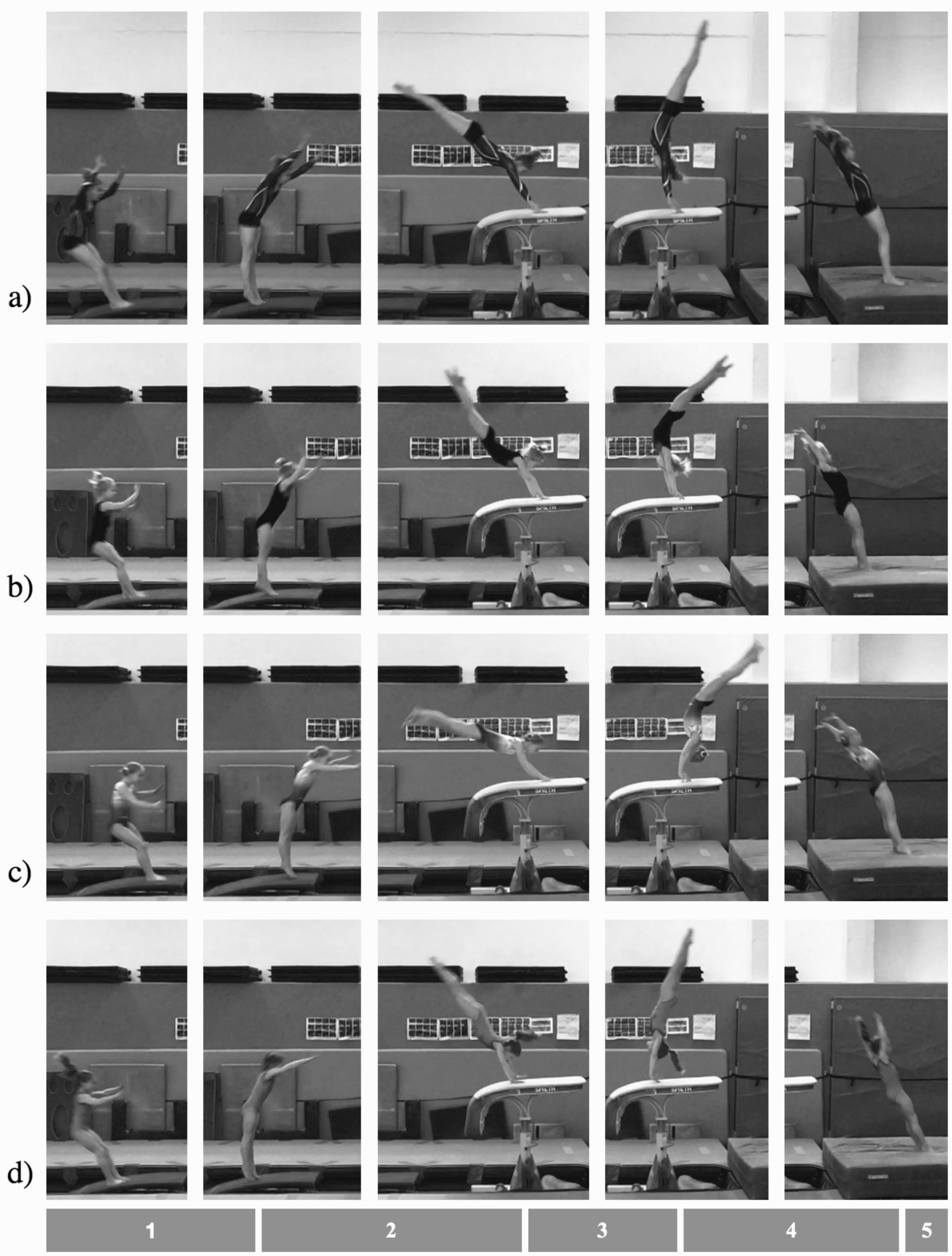

FIGURE 5.3: Illustration of the four handspring prototypes (see also Figures 5.1 and 5.2): a) Prototype #1, b) Prototype #2, c) Prototype #3, d) Prototype #4. *Note:* Each picture sequence shows one exemplary handspring trial of each prototype cluster. The letters *a)* to *d)* correspond to the four clusters in Figure 5.1. The numbers *1* to *5* correspond to the movement phases of the handspring (see Figure 5.2).

coefficient of variation of hip and shoulder joints for the four prototypes can be found in Figure 5.2 c) and Figure 5.2 d).

Prototype 2

In cluster #b, 15 trials were grouped together. A picture sequence of an exemplary trial can be seen in Figure 5.3 b). Exemplary time courses of hip and shoulder joints can be found in Figure 5.2 a) and Figure 5.2 b). A typical handspring trial from cluster #b comprised the following characteristics: 1) upright trunk orientation with shoulder angle slightly larger than 90° during touch-down on springboard, 2) inclined trunk, slightly flexed hip joint and open shoulder angle during take-off from the springboard, 3) slightly extended hip joint, slightly flexed shoulder joint and trunk orientation close to 45° during touch-down on the vaulting table, 4) considerably overarched back, stretched hip and shoulder joints during take-off from the vaulting table, and 5) straight back with slightly flexed hip and knee joints and open shoulder angle during touch-down on the landing mat. There is a small coefficient of variation for the hip angle and the knee angle over the whole movement (below 0.1). The shoulder angle showed a larger coefficient of variation (about 0.15) at the take-off phase, the beginning of the first flight phase, the end of the second flight phase and the landing phase and a coefficient of variation about 0.1 at the rest of the movement. Exemplary time courses of the coefficient of variation of hip and shoulder joints for the four prototypes can be found in Figure 5.2 c) and Figure 5.2 d).

Prototype 3

In cluster #c, 7 trials were grouped together. A picture sequence of an exemplary trial can be seen in Figure 5.3 c). Exemplary time courses of hip and shoulder joints can be found in Figure 5.2 a) and Figure 5.2 b). A typical handspring trial from cluster #c comprised the following characteristics: 1) upright trunk orientation with shoulder angle less than 90° during touch-down on springboard, 2) inclined trunk, slightly flexed hip joint and shoulder angle greater than or equal to 90° during take-off from the springboard, 3) slightly flexed hip joint, flexed shoulder angle and trunk orientation angle smaller than 45° during touch-down on the vaulting table, 4) considerably overarched back, stretched hip and flexed shoulder joints during take-off from the vaulting table, and 5) slightly overarched back with stretched hip and knee joints and open shoulder angle during touch-down on the landing mat. In terms of the variation of the movement for the different trials, there is a low coefficient of variation for all joint angles over the time course (0 - 0.1). Exemplary time courses of the coefficient of variation of hip and shoulder joints for the four prototypes can be found in Figure 5.2 c) and Figure 5.2 d).

Prototype 4

Finally, in cluster #d, 21 trials were grouped together. A picture sequence of an exemplary trial can be seen in Figure 5.3 d). Exemplary time courses of hip and shoulder joints can be found in Figure 5.2 a) and Figure 5.2 b). A typical handspring trial from cluster #d comprised the following characteristics: 1) upright trunk orientation with shoulder angle larger than 90° during touch-down on the springboard, 2) inclined trunk, slightly flexed hip joint and shoulder angle greater than 90° during take-off from the springboard, 3) slightly flexed hip joint, open shoulder angle and trunk orientation slightly greater than 45° during touch-down on the vaulting table, 4) straight back, trunk orientation about +10° from vertical, stretched hip and shoulder joints during take-off from the vaulting table, and 5) straight back with slightly flexed hip and knee joints and open shoulder angle during touch-down on the landing mat. In terms of the variation of the movement for the different prototypes, the knee angle shows a coefficient of variation of 0.2 at the take-off which decreased until almost zero at the end of the first flight phase. There was a small coefficient of variation for the hip angle(0 - 0.1) over the time-course. For the shoulder angle, the coefficient of variation was about 0.2 at the take-off phase and decreased to 0.1 during the first flight phase. Exemplary time courses of the coefficient of variation of hip and shoulder joints for the four prototypes can be found in Figure 5.2 c) and Figure 5.2 d).

Discussion

Most of the research concerning the kinematic analysis of gymnastics skills deals with selected variables. Because humans perceive movement in a holistic way, the goal of this study was to develop a method to analyze a front handspring on vault holistically. To gain insight into a complex motor skill like the handspring on vault, different prototypes of the movements should be quantitatively detected and the variant and invariant characteristics should be qualitatively investigated. The results of this study revealed that for near-expert gymnasts, four prototypical movement patterns could be identified. The four different prototypes can be differentiated by certain variant and invariant characteristics such as the time courses of the different joint angles and their coefficient of variation.

Concerning the assignments of the trials to one prototype, one can see that the trials from one person are not assigned consistently to the four prototypes, but rather to most of them. This highlights that the pattern of movement characteristics stays similar over a high number of trials, therefore indicating that there are structural similarities in space and time (Troje, 2002). When engaged in a goal-directed activity, like a handspring on vault, performers exhibit different coordination states. Those coordination states are not stable but demonstrate variability, leading to a set of movement options that could be described by a specific composition of biomechanical parameters (Latash et al., 2002). This inconsistency might explain why not all the trials of one gymnast are assigned to the same prototype. Variations in the movement patterns are carried out through an interaction of the body, as a complex biological system, with different constraints provided by the task, the environment or psychological factors, which lead to movement variability (Higgins, 1977).

Comparing the description of the four prototypes with the Code of Points (FIG, 2017), there are prototypes that meet the criteria for a high score and prototypes that might result in deductions. According to the Code of Points (FIG, 2017), there are deductions for a poor technique regarding the hips, the shoulder and the knees. Out of the four prototypes that have been identified, the movement patterns of prototype #a and prototype #d might most closely meet the criteria. They are characterized by extended knees and hips, and an open shoulder angle. The movement patterns of prototype #b and prototype #c might result in the worst scores.

There are limitations to this study, and three specific aspects should be highlighted. First, looking at the tree diagram of the cluster analysis, one might assume that the trials could also be distributed into two, five or even six clusters. When analyzing the pattern of the movement characteristics, it was revealed that by distributing the trials into two prototypes, a high number of structural features would have to be ignored, which could improve the description of the movement. On the other hand, taking five or six prototypes would not improve the description of the movement. These findings are in line with the results obtained from the elbow method, which looks at the percentage of variance explained as a function of the number of clusters.

Second, the study was conducted with near-experts at one point in time. For that reason, it is unclear whether the number of clusters and the allocation of the trials of one athlete to the different clusters are the same for top experts or novices, and how the allocation to the different clusters changes over time. One might assume that training leads to a change of the allocation of the movement execution to the different clusters. Either the movement patterns become restructured, which would be reflected in a more reliable allocation of different skill executions to one cluster, or the allocation of the skill executions of one athlete moves to a different cluster, which would be reflected in a less reliable allocation of different skill executions to one cluster.

Third, some effort should be made to study other gymnastic movements and their prototypical movement structures as well as how they appear in the variant and invariant features. This is relevant, particularly in gymnastics, because of the varying environmental constraints due to the different gymnastic apparatuses. The handspring is not only performed on vault, but it is also part of floor routines. The same fundamental movement has to be carried out in different ways, dependent on the features of the gymnastic apparatuses.

Furthermore, it would be interesting to investigate the relations between movement characteristics and the evaluation of the performance. It should be investigated whether the different prototypes are scored differently and how the movement characteristics, especially their variant and invariant features, find expression in the observers' gaze behavior, and when judging and evaluating the corresponding prototypes.

The current approach opens up interesting practical applications. With regard to gymnastics training, this study provides further evidence for the demand of individuality in training, in terms of an optimal organization of the complex functional movement system, in order to solve the movement task. By an adjustment of the skill execution of one athlete with the different prototypes, the skill level of the athlete could be easily determined and a specific training could be implemented. Depending on the similarity of the skill executions of one athlete to one specific prototype, different instructions in the training process might be beneficial.

Conclusion

Overall, the approach utilized in this study allows one to identify structural characteristics of movement patterns of a complex gymnastics skill. Therefore, this approach seems to be an appropriate and promising tool, not only for the analysis of gymnastics skills but also for a wide range of applications in various adjacent areas. The results open up practical both applications and further fruitful research questions.

Chapter 6

Publication II

The study described in this chapter has been published following peer review. Full reference details are: **Mack, M., Federbusch, S., Ferber, M., & Heinen, T. (in press). Movement prototypes and their relationship in the performance of a gymnastics floor routine. *Journal of Human Sport and Exercise.*** The author reserves the right that parts are slightly modified.

Abstract

For a better understanding of complex gymnastics performances, on the one hand, it is relevant to analyse isolated time discrete parameters, but on the other hand, it is also relevant to analyse the time-course of gymnastics skills and sequences holistically. Thus, the purpose of this study was to realize a holistic examination of a gymnastics floor routine (round off, back handspring, backward layout somersault) with an innovative approach of analysing time continuous data using a cluster analysis. Fifty-eight floor routine trials from ten female near-expert gymnasts were analysed on their movement kinematics. Time courses of six joint angles, together with the trunk orientation angle, were analysed by means of a hierarchical cluster analysis. In addition, the coefficients of variation were calculated. The results of this study revealed that for near-expert gymnasts, three to four prototypical movement patterns could be identified for each of the three skills (round off, back handspring, backward layout somersault). The different prototypical movement patterns can be differentiated by certain variant and invariant characteristics, such as the time courses of the different joint angles and their coefficients of variation. Statistically significant relationships were found between prototypes of the different gymnastics skills. In light of the training process in gymnastics, the study provides further evidence for strongly considering gymnasts' movement pattern as well as for focusing on particular movement characteristics rather than on particular gymnastics skills in regard to motor skill acquisition and optimization.

Introduction

Artistic Gymnastics involves very complex and technically challenging sequences of skills that require both maximum effort and a high level of functional ability, such as flexibility and coordination (Arkaev & Suchilin, 2004). Therefore, the movement-related components are linked to each other (Davids et al., 2014). For instance, the movement of the legs in a particular situation might influence the movement of the trunk, due to interacting joint torques (Enoka, 2002). Consequently, skills that are performed in sequence also interact with each other. Imagine a gymnast performing a round-off with a subsequent back somersault. The execution of the round-off (i.e., the variation in linear and angular momentum) influences the subsequent somersault (King & Yeadon, 2004). Thus, for a better understanding of gymnastics performance, the time course of the whole sequence of skills should be taken into account rather than collecting individual parameters (Troje, 2002). There are a large number of studies in the field of sports science and gymnastics, that investigate isolated biomechanical aspects of different gymnastics skills (Prassas et al., 2006). However, there is a lack of research concerning the investigation of gymnastics sequences in a holistic way, regarding the movement variability within one skill, but also relating to how this movement variability influences the relationship within a sequence of skills. Therefore, the goal of this study was to identify different prototypes and their variant and invariant characteristics, using an innovative method that allows the analysis of gymnastics skills in a holistic fashion. In a second step, the relationships between those prototypical movements, within a sequence of skills, were investigated.

Gymnastics skills, similar to other goal-directed activities, can be seen as complex systems. Complex systems are composed of many different variables that interact among themselves and with the environment as a whole . These interactions change, depending on the constraints provided by the task, the environment or the gymnast, thus making it impossible to completely preprogram their time-course in advance (Bernstein, 1967; Davids et al., 2014; Higgins, 1977). This, in turn, leads to a particular amount of variability that is reflected in distinguishable movement options that could be described by a specific structure of biomechanical parameters (Latash et al., 2002). Traditionally, movement variability has been treated as an aspect of human behavior that is not functional (Newell et al., 2006). Currently, through the introduction of nonlinear statistical models in the study of human movement systems, movement variability is considered to occupy a functional role in human motor behavior (Thompson & Stewart, 2002). It may be proposed that each skill can be defined behaviorally, by means of a unique set of interacting biomechanical properties (Newell, 1985; Newell et al., 2006).

For example, if the gymnastics skill of a somersault is performed by a certain number of gymnasts, a certain number of times, each of these performances is unique and can be described through a particular set of biomechanical properties (Schöllhorn et al., 2014). At the same time, however, some performances look more similar than others. This might imply that more similar performances of the same skill exhibit more similar biomechanical properties, and performances that vary significantly might be defined through different biomechanical properties. Furthermore, there might be some biomechanical properties, that exhibit less variability than others. Certain biomechanical properties are left free to vary, and in many cases, such variability could be viewed as a necessary and functional aspect of satisfying the task demands (Handford, 2006).

Biomechanical research has shown that these characteristics are not only descriptive, but also functional. For instance, the Tkatchev, a release and grasp skill on the high and uneven bars, provides a good example of the functional role of particular biomechanical properties of a complex skill. For instance, it could be shown that there was a greater angle of release for different versions of the Tkatchev, as the gymnasts extend the hip joint to reach the release angle (Irwin et al., 2011; Kerwin & Irwin, 2010). The modification of the hip joints can be seen as a functional aspect of changing the angle of release. Another functional role of biomechanical properties can be seen by the comparison of experts and novices. It is suggested that by freezing movements of joints and body segments, novices create a controlled solution to reduce the complexity of movements (Bernstein, 1967). In a later stage of learning, more joints and body segments are then 'freed'. It should be noted that individual constraints of different performers lead to different learning patterns in terms of the changes in performance and the biomechanical properties (Williams, Irwin, Kerwin, & Newell, 2012).

In gymnastics, especially during floor routines, several skills, such as a back handspring and a back somersault, are performed in very quick succession. Seeing gymnastics skills as complex systems, several skills performed in a sequence might not only interact among themselves and with the environment, but also with each other. Furthermore, a specific version of the back handspring should more likely lead to a specific version of the somersault. Individual constraints, such as the degree of expertise or the biomechanical properties, might be the cause of this variability. To perform a somersault after a back handspring, several requirements, such as a certain linear and angular impulse, have to be met during take-off from the floor, which should be prepared by the preceding skills. The question arises as to which biomechanical properties of the preceding skills are important for successful performance and how those biomechanical properties should be assessed.

Currently, a large amount of kinematic and kinetic data is available to describe human action. However, sports scientists usually identify, measure, and interpret selected variables, especially based on time-discrete amplitudes of selected variables (Federolf et al., 2012; Young & Reinkensmeyer, 2014). The emergence of dynamic systems theory as a viable framework for modeling the sensorimotor system stimulated a change of thinking in the assessment and investigation of human motor performance. For a valid understanding of human motor performance, the relationship of the biomechanical properties over the whole time-space continuum should be taken into account(Schöllhorn et al., 2014). As the pattern of the coordination states depends on the internal and external constraints, the time-space characteristics of the movement pattern provide insight into these constraints and the state of the dynamic system, not only at specific moments, but also in the movement pathway. Joints and body angles seem to be suitable variables to describe time-space characteristics of a gymnastics skill, because, on the one hand, they are continuous kinematic characteristics, while, on the other hand, they can be used to mathematically derive other time-discrete variables (Prassas et al., 2003).

Quantitative technique analysis seems inappropriate for explaining the characteristics of the whole skill, but multivariate statistical methods, such as cluster analysis or principal component analysis, may be able to overcome the aforementioned limitation (Davids et al., 2014; Lees, 2002). Both approaches use relative similarities or proximities of variables as a theoretical statistical basis that matches the idea of movement as a unique set of interacting biomechanical properties. A commonly used measure for mathematical comparisons is the Euclidean distance, which represents the mathematical distance between two objects (Bauer & Schöllhorn, 1997; Schöllhorn et al., 2014). Cluster analysis quantitatively sorts these Euclidean distances, thereby aiming to find groups of objects with a high degree of structural similarity to each other, which can be visualized in a tree diagram (Everitt & Dunn, 2001). Given the natural variation of objects in relation to their analyzed qualities, the different clusters contain a certain degree of variability (Troje, 2002).

It can be stated that for a better understanding of complex gymnastics performances, it is relevant not only to analyze isolated parameters, but also to analyze gymnastics skills in terms of the interacting biomechanical properties over their entire time-course. Therefore, the goal of this study was to identify different prototypes and their variant and invariant characteristics by means of an innovative method that allows the analysis of gymnastics skills in a holistic fashion. In a second step, the relationships between those prototypical movements in a sequence of skills were investigated.

In the first step, the time course of gymnasts' joint angles, in separate trials of a sequence of three specific gymnastics skills (round off, back handspring, somersault), were mathematically analyzed by means of a cluster analysis. It was hypothesized that some trials are more similar than others. The cluster analysis should reveal patterns of similarity, leading to a particular number of distinguishable clusters (i.e. prototypes). In the second step, the relationships between those prototypes were investigated. In the third step, the variant and invariant characteristics were evaluated qualitatively by analyzing the time courses of the joint angles in relation to the different prototypes and the different movement phases. Because some biomechanical properties of movement organization are more variable than others, it is hypothesized that the prototypes will differ in their variant and invariant characteristics in specific phases of movement.

Material and Methods

Participants

Ten female gymnasts (age: $M = 11.50$ years, $SD = 1.43$ years) with an average training time of 26 hours per week participated in this study. They were able to perform the experimental task of this study with a high degree of consistency in training and competition (floor routine; see Motor task section). All participants were given verbal information concerning the nature and purpose of the study, and they gave their verbal consent. The gymnastics movements realized for data collection were the same as those performed by the gymnasts during their regular training. The data collection process was nonreactive. There was no manipulation of the participants in the study; they had no time pressure and were fully informed regarding the goal and the procedure of the study. The study was conducted in compliance with the Helsinki Declaration for Human Research and the International Principles governing research on humans, and was in line with the ethical guidelines of the local ethics committee.

Motor Task

The motor task was a gymnastics sequence on the floor, containing the following gymnastics skills: (1) round-off, (2) back handspring and (3) back layout somersault. The three skills were performed consecutively on a spring floor that was set up according to the competition guidelines of the International Gymnastics Federation for women's artistic gymnastics (FIG, 2017). Gymnasts were instructed to perform the skills as they would do in a regular competition. The moment at which one skill ends and the consecutive skill starts is defined as the point of time when the gymnast touches the floor in the second flight phase, to transit to the take-off phase of the consecutive skill. This

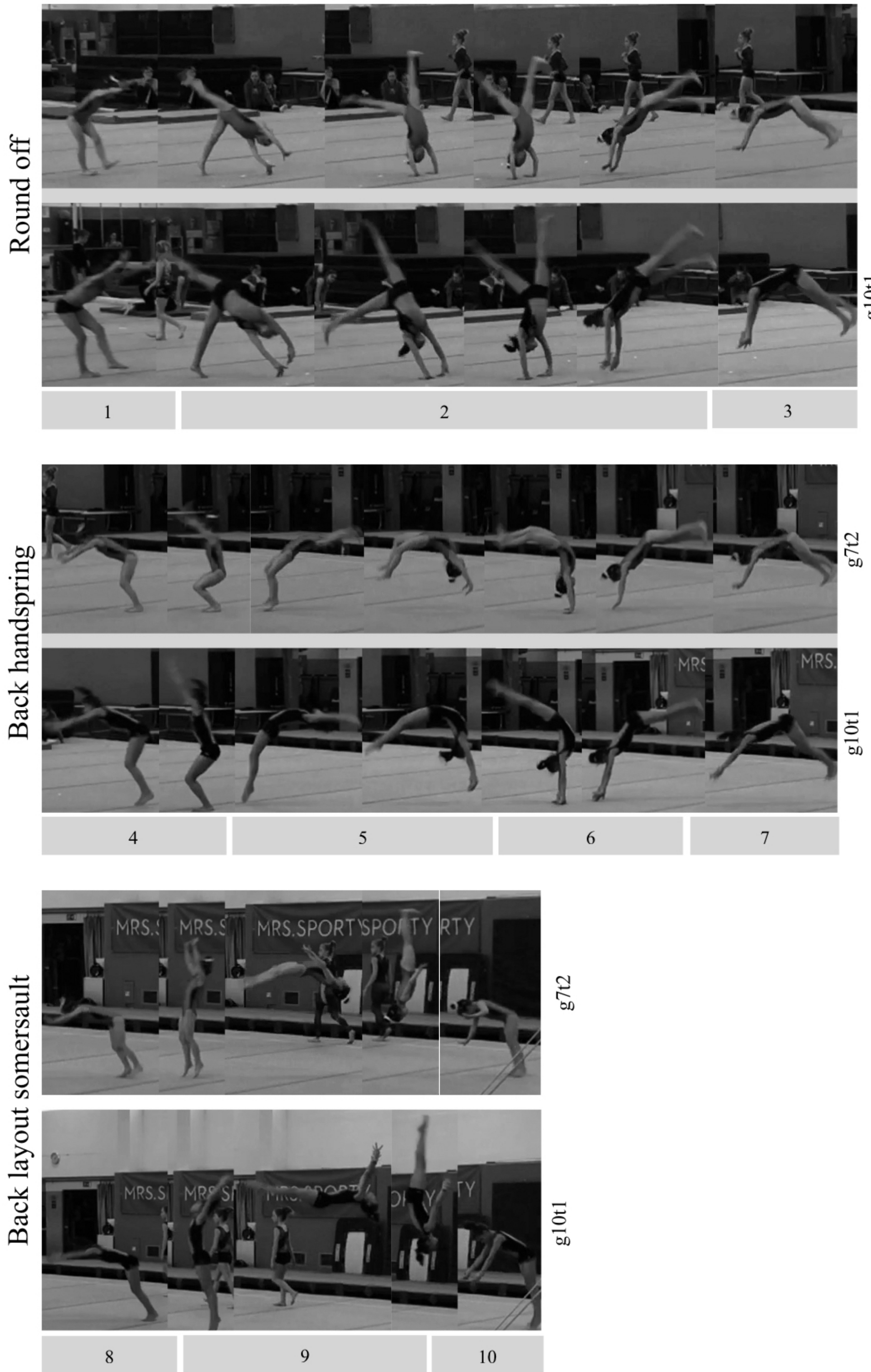

FIGURE 6.1: Illustration of two prototypical trials of the floor routine. *Note*: The numbers "1" to "10" correspond to the movement phases of the three skills: round-off g721 = cluster #2, g10t1 = cluster #1; back handspring g7t2 = cluster #3, g10t1 = cluster #1; backward layout somersault g7t2 = cluster #4, g10t1 = cluster #1.

leads to the following description of the three skills and the corresponding classification of the movement phases of the three skills (Figure 6.1):

Round-off: The round-off contains the following movement phases: (1) first flight phase, (2) support phase and (3) second flight phase. It begins with an initial hop that is performed after a short run-up. The gymnast raises her arms up and forwards. The body rotates forward while the leading leg steps forward, thereby generating initial angular momentum. The leading hand is placed on the floor, in front of the gymnast. If the leading hand is the right (left) hand, the second hand is placed further to the right (left) of the center line, facing backward. The bent leg (right) pushes and the rear leg swings overhead, thereby further translating the initial kinetic energy into whole-body rotation. An accentuated push from the shoulder supports the lift. The turn is made during the hand placement and through the pushing stage from the hands to the feet. By creating a concave shape of the back, the feet touch the floor, and the shoulders rise rapidly with the arms in front of the shoulders (Karácsony & Čuk, 2005; Turoff, 1991).

Back handspring: The back handspring contains the following movement phases: (4) take-off phase, (5) first flight phase, (6) support phase and (7) second flight phase. The skill starts with a jump backward from a standing position. The gymnast moves her arms overhead and extends her hips, thereby engaging an arched body posture. The hands are turned so that the fingers face each other. As the hands come into contact with the mat during the support phase, the elbows may be slightly flexed. Once the body passes over the hands, the legs are snapped down forcefully, while the hands are pushed down hard onto the mat. This movement supports taking off from the floor at the end of the support phase. During landing, the body is in a slightly inclined position, with the arms in front, and the knees and hips flexed slightly (Karácsony & Čuk, 2005; Turoff, 1991).

Backward layout somersault: The backward layout somersault contains the following movement phases: (8) take-off phase, (9) flight phase and (10) landing phase. This skill usually begins from a take-off position at the end of the back handspring. The arms swing upward and slightly backward with the shoulders above the feet. The center of gravity should be on or in front of the gravitational vertical. The hips are pushed forwards and upwards during take-off. During the flight, the body is tightened, and the head is held in a neutral position. The landing is prepared after the shoulders have passed the vertical line. During landing, the knees are slightly bent to absorb the kinetic energy (Karácsony & Čuk, 2005; Turoff, 1991).

Movement Analysis

To record the performance of the gymnasts, a digital video camera (240 Hz, 1920 x 1080 pixel) was placed at a distance of approximately 15 meters from the place on the floor where the routine was to be performed. The gymnasts' performances were videotaped orthogonally to the direction of movement. With the movement analysis system Simi Motion®, the recorded video sequences were digitized for each frame. The digitized horizontal and vertical coordinates of ten points (body landmarks) defined an eight-segment model of the human body (Enoka, 2002). Each body landmark was represented by a two-dimensional time series $[x_j(t);\ y_j(t)]$ with $j = 1, 2, 3, ..., j$ (t = time, j = frame number). For each video sequence, a digital filter built into the software was applied for data smoothing. The time series of the ten body landmarks were cut into three parts, with each of those parts representing one trial of one particular gymnastics skill of the floor routine. For each trial, the time series of each body landmark was time-normalized and rescaled to the interval [0; 1000]. Joint angles of six joints and the trunk orientation angle were calculated from the time-normalized position data of the body landmarks for all handspring trials (Jaitner, Mendoza, & Schöllhorn, 2001). The calculated joint angles (knee, hip, and shoulder of the left and right body sides), reflecting flexion and extension movements, were calculated with regard to the frontal horizontal body axis (Behnke, 2001).

Procedure

The study was conducted in three phases. In the first phase, the gymnast arrived at the gymnasium. She was instructed regarding the general procedure of the study and was told that she was taking part in a study on kinematic analysis of a gymnastics floor routine. After the gymnast gave her informed consent, she was given a warm-up period of approximately 20 minutes and one optional familiarization trial. In the second phase, the gymnast performed ten trials of the motor task. There was no time pressure, and she was allowed to take breaks as requested. In the third phase, after completing the ten trials, the gymnast was released into an individual cool-down period.

Data Processing and Analysis

For further data processing and analysis, the free statistics software R (R Core Team, 2017) was used. First, the prototypes of the three skills of the gymnastics routine on the floor (round-off, back handspring and backward layout somersault) were determined using a hierarchical cluster analysis (Mack, Hennig, & Heinen, 2018). For each time course of joint angles (knee, hip and shoulder of the left and right body sides) and the trunk orientation angle (right and left body sides), Euclidean distances were calculated between each pair of the same skill trials (see equation 1: x and y denote a corresponding joint angle between a pair of two trials of the same gymnastics skill, and i denotes a point in the rescaled time interval [0;1000]). An exactly identical course of two skills of the same course would have a value of zero, and the more dissimilar two trials were, the larger the resulting value. Therefore, a value of zero would have indicated an exactly identical course of two skill trials, whereas the larger the resulting value, the more dissimilar the two trials were.

$$d(x, y) = \sqrt{\sum\nolimits_{i=1}^{n} (x_i - y_i)^2} \tag{6.1}$$

A distance matrix resulting from of the calculated Euclidean distance values indicates the similarity between each pair of two trials of the same gymnastics skill. With a hierarchical cluster analysis using Ward's hierarchical clustering method (Ward, 1963), all trials of the same gymnastics skill were evaluated quantitatively, according to their similarity. Ward's method is an agglomerative clustering method that is based on a classical sum-of-squares criterion, producing groups that minimize within-group dispersion at each fusion step (Murtagh & Legendre, 2014). The resulting two-dimensional tree diagram emerged through the fusions or divisions made at each stage of the analysis, illustrating the classification of the trials. The number of clusters was specified by interpreting the scree plot in terms of the elbow criterion (Everitt & Dunn, 2001). In the second step, the relationship between the prototypical movements in a sequence of skills was investigated with Cramer's V correlation method. In the third step, for each cluster, the time courses of the joint angles and the trunk orientation angle were averaged over the corresponding trials, in order to characterize each of the prototypes. To obtain the relative extent of variability of a particular prototype in the passage of time, the time courses of the coefficient of variation were calculated for all joint angles, as well as for the trunk orientation angle of each cluster (Stergiou, 2004).

Results

Structuring of Gymnastics Skills

Figure 6.4 presents the results of the hierarchical cluster for the three gymnastics skills (round-off, back handspring, backward layout somersault) performed by the gymnasts.

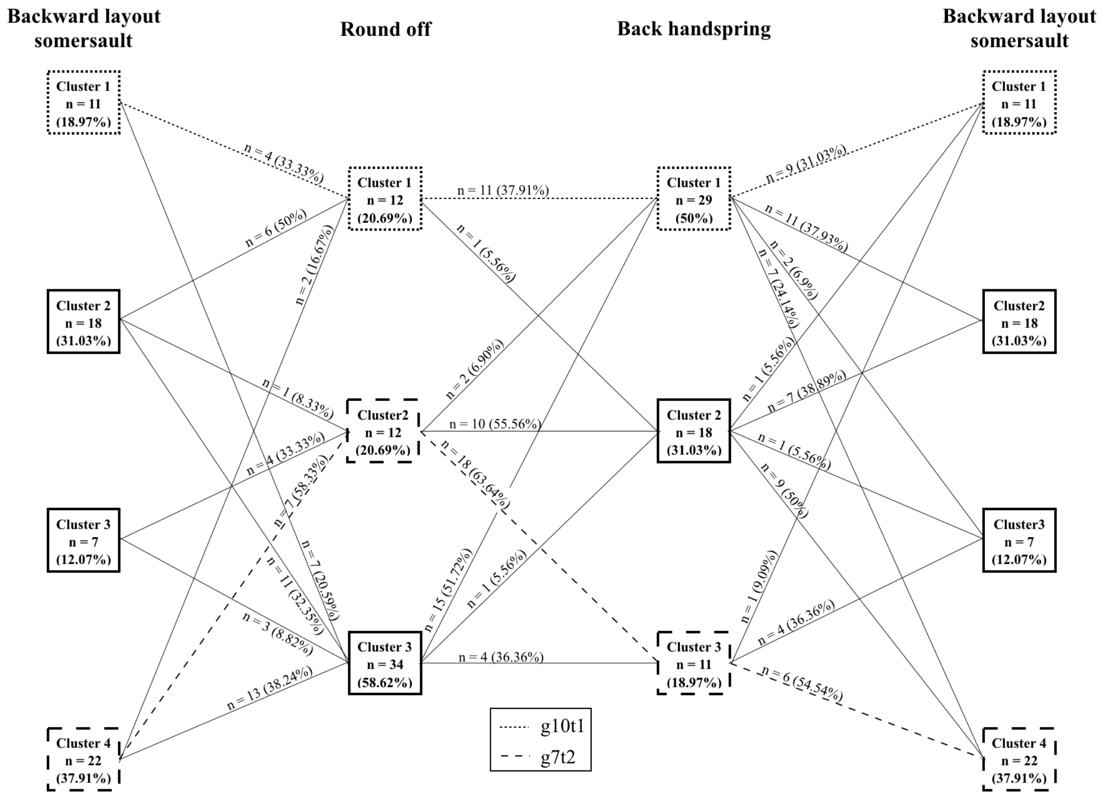

Figure 6.2: Results of the analysis of the relationships between the prototypes / clusters. *Note*: The lines illustrate the amount and percentage of the trials of a cluster that are related to the following cluster.

The clusters could be distinguished from each other following the inspection of the scree plot of the cluster analysis. Each of the clusters comprised trials of the respective skill that were more similar to trials within a particular cluster, but were more dissimilar to trials in the other clusters. Therefore, each cluster characterized a particular prototype of the respective skill, with specific time-space characteristics of the movement.

Relationship between the Skill Prototypes

To investigate the relationship between the three gymnastics skills in a sequence, Cramer's V correlations between the assignments of the trials to the different clusters for the different gymnastics skills were conducted. The assignment of a trial to a specific cluster in terms of the round-off significantly correlated with the assignment of a trial to a specific cluster in terms of the back handspring, $\chi^2(4, N = 58) = 47.986$, $p < .001$, as well as in terms of the backward stretched somersault, $\chi^2(6, N = 58) = 16.021$, $p = .014$. The assignment of a trial to a specific cluster in terms of the back handspring significantly correlated with the assignment of a trial to a specific cluster in terms of the backward stretched somersault, $\chi^2(4, N = 58) = 18.317$, $p < .005$. To determine

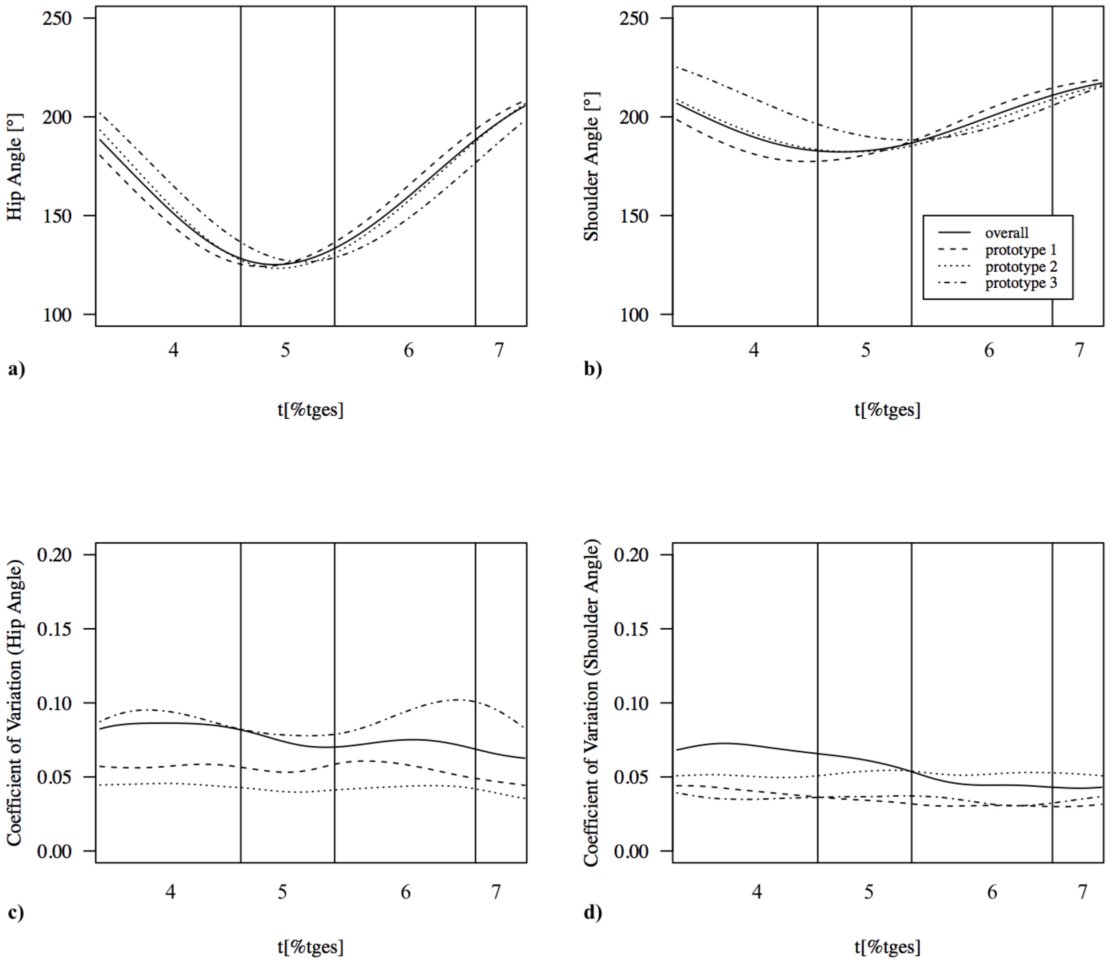

FIGURE 6.3: Illustration of time-normalized angle-time plots for the prototypical courses of the right hip angle (a) and the right shoulder angle (b), as well as time courses of the corresponding coefficients of variation for the different prototypes (c, d) for the back handspring. *Note*: 4 = take-off phase, 5 = first flight phase, 6 = support phase, 7 = second flight phase.

how the assignment of the gymnastics skills in a movement sequence depends on the foregoing skills, the percentage of the trials of a cluster, found again in the different clusters of the following skill, was calculated. The results are presented in Figure 6.2 . Each of the clusters thus comprised trials of the respective movement sequences that were assigned to different clusters in the preceding or following gymnastics skill.

Description of the Prototypes

Each cluster characterized a particular movement prototype within the sample of all analyzed movement trials. Some of these clusters were related and some were not. Therefore, a closer look should be taken at the movement characteristics of the clusters of two particular trials (Trial 2 of Gymnast 7 and Trial 1 of Gymnast 10) which were assigned to different clusters in the three gymnastics skills. A picture sequence of an

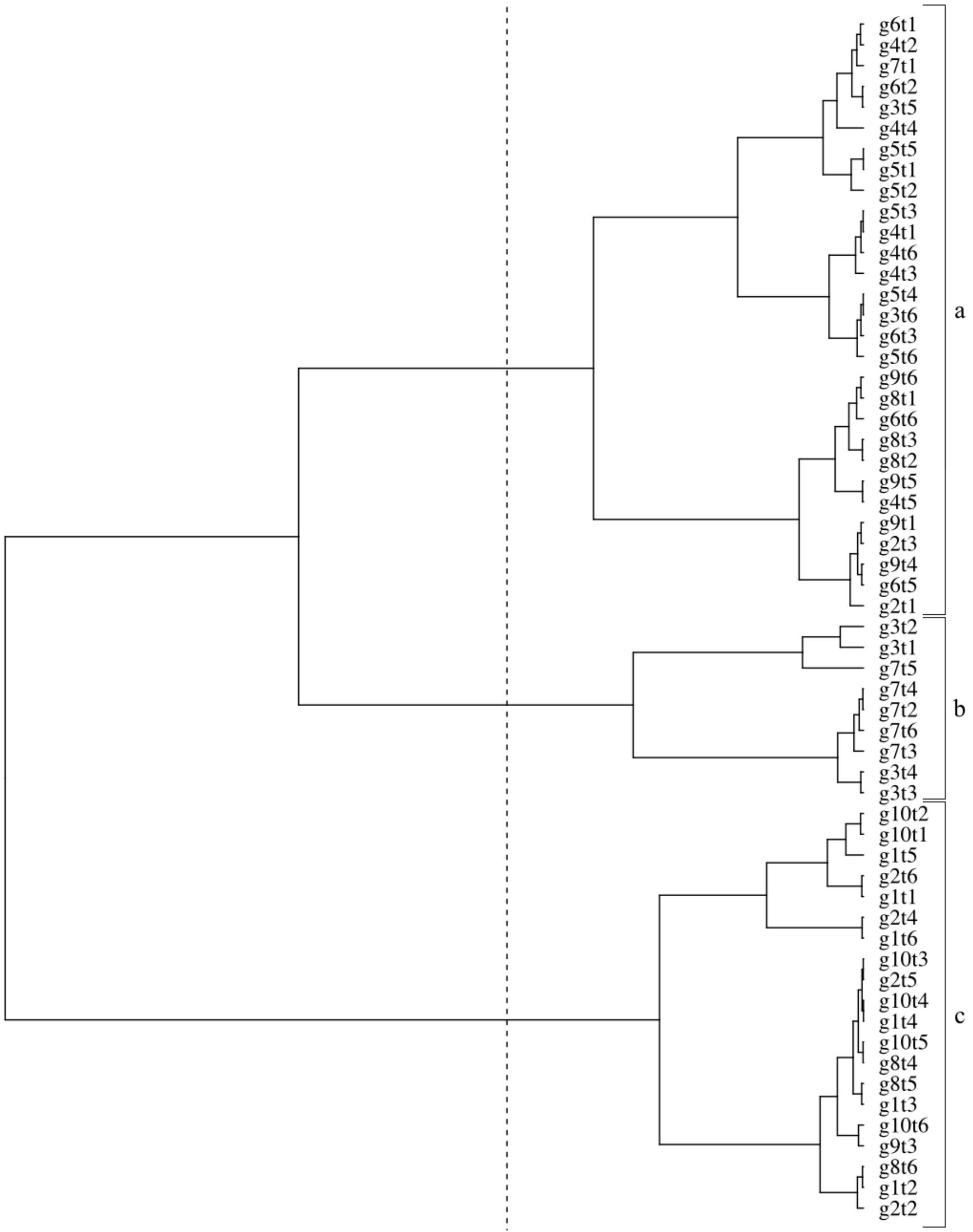

Figure 6.4: Tree diagram resulting from a cluster analysis using Wards' clustering algorithm. *Note*: Horizontal lines indicate the level of the distance at which the respective backward layout somersault trials are grouped into one cluster. The dashed line represents the Euclidean distance below which the clusters are identified. The letters "a)" to "c)" correspond to the three clusters containing the different prototypical movement patterns of the gymnastics floor routine. g1t1 to g10t6 represent the analysed backward layout somersault trials.

exemplary trial can be seen in Figure 6.1. Exemplary time courses of hip and shoulder joints, as well as their coefficients of variation for the back handsprings, can be found in Figure 6.3 .

Round off: Regarding the round off, Trial 2 of Gymnast 7 was assigned to Cluster #2 and Trial 1 of Gymnast 10 was assigned to Cluster #1. In Cluster #2, twelve trials were grouped together. A typical round-off trial from Cluster #2 comprised the following characteristics: (1) Open shoulder angle and flexed left hip joint as well as open right hip joint during the first flight phase. (2) Stretched elbow and shoulder joints, as well as stretched knee angles, and a stretched back during the support phase. (3) Flexed shoulder and hip joint as well as stretched knee and elbow joints during the second flight phase. There was a rather small coefficient of variation for the right hip and the right shoulder joints over the time course (cv = 0.03 - 0.07), especially at the beginning and the end of the round-off.

In Cluster #1, twelve trials were grouped together. A typical round off trial from Cluster #1 comprised the following characteristics: (1) Open shoulder angle and flexed left hip joint as well as open right hip joint during the flight phase. (2) Slightly flexed elbow and shoulder joints as well as stretched knee angles during the support phase. (3) Flexed shoulder and hip joint as well as stretched knee and elbow joints during the flight phase. There was a rather small coefficient of variation for the right hip and the right shoulder joints over the time course (cv = 0 - 0.1), decreasing to the end of the round-off.

Back handspring: Regarding the back handspring, Trial 2 of Gymnast 7 was assigned to Cluster #3 and Trial 1 of Gymnast 10 was assigned to Cluster #1. In Cluster #3, eleven trials were grouped together. A typical back handspring trial from Cluster #1 comprised the following characteristics: (4) Slightly flexed hip and knee joints during the take-off phase; (5) over-arched back, slightly bent knee and slightly flexed shoulder angle during the first flight phase; (6) slightly bent knees, an open shoulder angle and a hip angle that started bending during the support phase; (7) open shoulder angle and flexed hips during the second flight phase. There was a rather small and steady coefficient of variation of approximately cv = 0.05 for the right shoulder joint over the time course and a coefficient of variation between cv = 0.05 and cv = 0.1 for the right hip joint, increasing at the beginning and the end of the back handspring.

In Cluster #1, 29 trials were grouped together. A typical back handspring trial from Cluster #4 comprised the following characteristics: (4) Slightly flexed hip and knee joints during the take-off phase; (5) over-arched back, slightly stretched knee and open shoulder angle during the first flight phase; (6) stretched knees, an open shoulder angle and a hip angle that started bending during the support phase; (7) open shoulder angle and flexed hips during the second flight phase. There was a rather small and steady coefficient of variation of approximately cv = 0.05 for the right shoulder joint and the right hip joint over the time course.

Backward layout somersault: Regarding the backward layout somersault, Trial 2 of Gymnast 7 was assigned to Cluster #4 and Trial 1 of Gymnast 10 was assigned to Cluster #1. In Cluster #4, eleven trials were grouped together. A typical backward layout somersault trial from Cluster #4 comprised the following characteristics: (8) Slightly bent knees, shoulder angle slightly above 90° and a trunk orientation of approximately 45° during the take-off phase; (9) open hip angle, stretched knees and a shoulder angle of approximately 45° during the flight phase; (10) slightly bent knees, slightly closed hips and a shoulder angle of approximately 45° during the landing phase. There was a coefficient of variation between cv = 0.03 and cv = 0.11 for the shoulder joint, decreasing during the flight phase and a coefficient of variation between cv = 0.05 and cv = 0.15 decreasing during the backward layout somersault.

In Cluster #1, eleven trials were grouped together. A typical backward layout somersault trial from cluster #1 comprised the following characteristics: (8) Slightly bent knees, shoulder angle slightly above 90° and a trunk orientation of approximately 45° during the landing take-off phase; (9) open hip angle, an over-arched back, stretched knees and a closed shoulder angle during the flight phase; and (10) slightly bent knees, slightly closed hips and a shoulder angle approximately 45 during the landing phase. There was a coefficient of variation between cv = 0.05 and cv = 0.11 for the shoulder joint, decreasing during the flight phase, and a steady coefficient of variation of approximately cv = 0.05 over the time course of the backward layout somersault.

Discussion and Conclusion

For a better understanding of complex gymnastics performances, it is relevant not only to analyze isolated parameters but also to analyze gymnastics skills in terms of the interacting biomechanical properties over their entire time course. Thus, the purpose of this study was to realize a holistic examination of gymnastics skills with an explorative approach of analyzing time continuous data and using the detected classification to exploratively investigate how the particular skills correlated with each other in a sequence of gymnastics skills.

The results of this study revealed that for near-expert gymnasts, three to four prototypical movement patterns could be identified for each of the three skills (round-off, back handspring, backward layout somersault) as well as the complete gymnastics floor routine. The different prototypes can be differentiated by means of certain variant and invariant characteristics, such as the time courses of the different joint angles and their coefficient of variation. There is a statistically significant relationship between some prototypes of the different gymnastics skills.

Concerning the assignments of the trials to the prototypes, it is obvious that the trials from one person are not assigned exhaustively to one specific prototype for all three gymnastics skills. Complex systems such as a round-off, a back handspring or a backward layout somersault are composed of many different variables that interact among themselves, and with the environment as a whole. These interactions change, depending on the constraints provided by the task and the environment or the gymnast him or herself, and without being previously planned (Bernstein, 1967; Davids et al., 2014; Higgins, 1977). This leads to the variability that is reflected in distinguishable movement options that could be described by a specific structure of biomechanical parameters, and it might explain why not all trials of one gymnast are assigned to one prototype (Latash et al., 2002).

Points (FIG, 2017), there are prototypes that might receive high scores, and there are prototypes that might receive deductions, and thereby lower scores. According to the Code of Points (FIG, 2017), there are deductions for poor technique regarding the movements of the hips, the shoulders, and the knees. Out of the two exemplary movement sequences, the movement pattern of Trial 1 of Gymnast 10 might obtain a better overall score for the three gymnastics skills than the movement pattern of Trial 2 of Gymnast 7. In summary, the differences for the round-off were especially apparent in movements of the shoulder and the elbow joints for all movement phases. Trial 1 of Gymnast 10 showed more stretched elbow and shoulder joints than Trial 2 of Gymnast 7. For the back handspring, differences were found, especially in the (5) first flight phase and the (6) support phase, in terms of more flexed shoulder and knee joints for Trial 2 of Gymnast 7 than for Trial 1 of Gymnast 10. For the last skill, the backward layout somersault, Trial 1 of Gymnast 10 and Trial 2 of Gymnast 7 differed in terms of a more over-arched back and closer shoulder joints for Trial 2 of Gymnast 7, compared to Trial 1 of Gymnast 10, during the (9) flight phase.

Looking at the manifestation of the particular movement characteristics, for example, the time-space pattern of the knee, shoulder or hip joints, one can see that particular movement characteristics of the different prototypes of one of the three gymnastics skills are found again in the different prototypes of one of the other gymnastics skills. Both the extended knee and hip angle and the open shoulder angle are characteristics that are consistent throughout the entire floor routine in Trial 1 of Gymnast 10 and might explain the revealed correlations.

From the perspective of the dynamic system theory, the fact that there are different prototypes that emerge under a different set of constraints leads to some considerations concerning the potential role of the prototypes (i.e., coordination states) in terms of motor learning. It may not be so much a matter of physical preparation that Gymnast 7 performs worse than Gymnast 10, but it could be more a matter of motor coordination, i.e., exploring degrees of freedom. By 'freezing' movements of joints and body segments, novices create a controlled solution to reduce the complexity of movements. In a later stage of learning more joints and body segments are then 'freed' (Bernstein, 1967). Through learning, the entire layout of the coordination dynamics changes. The prototypes might reveal the qualitative changes of the movement and may also be the point at which behavior bifurcates or the gymnasts change their learning phases.

This leads to a further aspect. It is unclear whether the number of the clusters of the different gymnastics skills and the distribution of the trials of one athlete to the different clusters are the same for all skill levels, from novices to top experts, or how the distribution changes through training. One might assume that training leads to a change in the number of clusters and in the distribution of the trials to the different clusters. This is in line with the results of the functional role of biomechanical properties when comparing experts and novices. It is suggested that by 'freezing' movements of joints and body segments, novices create a controlled solution to reduce the complexity of movements. In a later stage of learning, more joints and body segments are then explored (Bernstein, 1967). It should be noted that individual constraints of different performers lead to different learning patterns in terms of the changes in performance and biomechanical properties (Williams et al., 2012).

A second aspect is related to the question of finding the optimum point in time to proceed to learning more complex versions of a particular skill, for example, a double somersault instead of a single somersault. The perfect point in time might be related to a particular skill performance, which is similar to a particular prototype.

Those considerations in terms of the current approach open up practical implications. With regard to gymnastics training, it can be stated that individuality in training in terms of an optimal organization of the complex functional movement system is an important factor in learning and improving gymnastics skills. The skill execution of one athlete can be categorized by an adjustment of the performance with the different prototype, and beneficial instructions for the training process can thereby be selected. This study provides further evidence that specific biomechanical properties, for example, an open hip angle, were reflected, not only in a single gymnastics movement, but also in the different gymnastics movements. For the training process, this might imply that it would be useful to focus on the biomechanical property that should be improved, and not on the particular gymnastics skills. Furthermore, the approach might be used as a diagnostic tool to ascertain stations in the learning process, at which to start learning and training new skills or more complex versions of a particular skill.

There are limitations of this study, and two specific aspects should be highlighted. First, each of the three gymnastics skills was structured into three to four clusters. One might assume that the trials could also be distributed in more or fewer clusters. However, by distributing the trials into fewer clusters, a high number of structural features might be ignored, and by distributing the trials into more clusters, the description of the movement might not be improved. These findings are in line with the results obtained from the elbow method, which looks at the percentage of variance explained as a function of the number of clusters, and by the fact that the gymnasts were near-experts concerning the realized gymnastics skills in the sequence. In further studies, the role of gymnastics expertise or motor learning processes in the number of clusters of the different gymnastics skills and the distribution of the trials of one athlete to the different clusters should be investigated. This might, on the one hand, lead to more specific practical implications in terms of the biomechanical properties that are important for gaining the skills required to deliver an expert performance. On the other hand, this might lead to better insight regarding the best point in the training process at which to learn new skills, without training too long for a particular skill and losing the performance level of a particular age group, or training for too short a time, leading to a lack of the basic performance level.

Furthermore, it would be interesting to investigate whether the different prototypes are scored differently, and how the movement characteristic finds expression in the gaze behavior of the observers when judging and evaluating the corresponding prototypes. For floor routines containing more than one gymnastics skill, it should be determined whether there is a primacy or recency effect, i.e., the first or the last gymnastics skill has the highest influence on the judging behavior, because of the higher likelihood of it being remembered. Another interesting point is the questions of whether changing the movement characteristics of the beginning skills in a sequence of movements influences the movement characteristics of the preceding skills.

Overall, the approach utilized in this study allows one to identify structural characteristics of movement patterns, not only of a complex skill but also of their relationships in a sequence of those complex skills. This approach might work well for analyzing movement in various adjacent areas, and it can easily be adjusted to a wide range of applications and research questions. The results open up both practical applications and further research questions.

Chapter 7

Publication III

The study described in this chapter has been submitted following peer review. Full reference details are: **Mack, M., Schmidt, M., & Heinen, T. (submitted). Mental representation of movement quality of complex skills in gymnastics.** The author reserves the right that parts are slightly modified.

Abstract

The study aimed to investigate the relationship between the perceived movement quality of a gymnastics skill and its kinematics pattern, as well as the influence of expertise. Thirty participants with different levels of gymnastics expertise ($n = 10$ visual experts, $n = 10$ motor experts and $n = 10$ novices) were recruited for the study. They were instructed to compare the movement quality of eleven video sequences, showing different handstand - back handspring performances. To extract the kinematics, the performances were digitized. By means of an ongoing cluster analysis, both the kinematics pattern and the pattern of the perceived movement quality of the skills were determined for each experimental group. The results of the cluster analysis of the different experimental groups were analyzed and compared. Differences were found in terms of the expertise, regarding the pattern of the perceived movement quality. Novices thereby showed a higher variance in their judgment than the visual and motor experts. The pattern of the perceived movement quality was not correlated with the holistic kinematic pattern of judged skills. These results suggest perceptual and cognitive differences in the participants, due to differences in their previous visual and motor experiences.

Introduction

As in other gymnastics disciplines or sports, such as synchronized swimming or figure skating, the determination of sporting performance in artistic gymnastics relies on the judge's estimation concerning the movement quality of the presented skills (Plessner & Haar, 2006). Therefore, perceiving and evaluating skills plays a crucial role for judges, coaches, and the athletes themselves. While judges allocate points in a competition to identify the best athlete, coaches and athletes try to find sources of error to improve the movement. For both sides, the main idea is to determine the movement quality of a particular skill or routine and to discriminate between bad and good movement quality as precisely as possible (Mercier & Heiniger, 2018). The aim of this study was therefore to address the representation of movement quality. More specifically, both the relationship between the perceived movement quality of a gymnastics skill, and its kinematics pattern, and the influence of expertise were explored.

In gymnastics, the performance and thereby movement quality-determining aspects (i.e. straight body), by which the skills should be evaluated, are described in the code of points (Arkaev & Suchilin, 2004; FIG, 2017). Every skill is evaluated with regard to its execution and difficulty. The difficulty panel assigns each skill a particular score that reflects the technical demands of the skill, whereby the quality-determining score is assigned by the execution panel. Thus, to obtain a performance score, the presented skill has to be compared with an optimal model (Ste-Marie, 1999).

Gymnastics skills can be seen as a complex system, consisting of many different components that interact among themselves and with the environment (Bernstein, 1967; Davids et al., 2014; Higgins, 1977). This interaction leads to a high amount of movement variability reflected in distinguishable movement options. Therefore, each skill can be described by a specific structure of biomechanical parameters that are more or less similar to each other (Latash et al., 2002, Newell, 1985; Oullier, Marin, Stoffregen, Bootsma, & Bardy, 2006).

It is suggested that movements are organized and stored in the memory as a cognitive representation. These cognitive movement representations are based on the temporal and spatial kinematic features of movements acquired during training (Schack, 2010) and vary according to expertise. According to Schack's architecture model (Schack, 2004; Schack & Ritter, 2009), mental representations consist of functional and sensory features. The functional features refer to movement goals utilizing the anticipated effects of the movement. The sensory features that were integrated into the movement through cognitive chunking refer to the perceptual effects of the movement.

There is an increasing amount of research addressing the mental representation of movements in terms of their functional states and sensory events, as well as the presence of kinematic information (Schack, 2003). In sports such as artistic gymnastics or dancing, not only are the anticipated effects of a movement of importance, but it is also essential to evaluate the movement quality of a skill as a whole. From that point of view, not only the representation of the successive detailed action steps is relevant, but also the representation of the relationship of the kinematics variables over the whole time and space during which the movement is performed. To evaluate a gymnastics skill, all biomechanical information has to be summed up and integrated into an evaluation.

It is assumed that when perceiving movement, kinematic information, which is distributed over the whole time and space of the action, is taken into account (Troje, 2002). The perceptual process may be moderated by aspects such as observers' expectations or expertise. Different levels of expertise could be found concerning the perception of biological motion in daily life, where the movement recognition seems to be influenced by the observers' familiarity with the observed action (Aglioti, Cesari, Romani, & Urgesi, 2008; Loula, Prasad, Harber, & Shiffrar, 2005) and by the perception of different complex dance movements (Bläsing & Schack, 2012; Calvo-Merino, Ehrenberg, Leung, & Haggard, 2010).

Due to the results of research concerning motor-visual expertise and the underlying theoretical framework of the common coding theory (for a broader view of this model, view Prinz, 1997; Hommel et al., 2001), it could be assumed that in gymnastics, differences might exist not only between experts and novices but also between visual and motor experts, concerning the perception and information processing of sports skills (Kim, Park, & Schack, 2019). By means of the specific and comprehensive qualifications regarding the quality-determining criteria, coaches and judges should gain precise knowledge about the differences of the skills in terms of their kinematic features and their matched quality scores. Therefore, it could be assumed that visual experts, such as coaches or judges who have a more detailed and precise representation of this optimum execution model, are consequently better to observe movement errors than motor experts (athletes) or novices, and might determine a more accurate estimation of the movement quality (Land et al., 2013; Ste-Marie, 2001).

The question remains as to whether movements with a more similar structure of biomechanical parameters were seen as more similar to movements with a less similar structure, and whether a similar structure of movement parameters leads to a similar judgment of movement quality. While there is research that deals with searching for the path between biomechanical aspects of movement and the functional elements of movement organization, there is a missing link between biomechanics and the perceived movement quality. In sports for which the judgment relies on subjective perceptual analysis, the relationship between the biomechanics and the perceived movement quality is poorly understood.

The aim of this study was to address the representation of movement quality. Therefore, two research questions were investigated. First, to identify the relationship between the kinematic structure of complex gymnastics skills and the mental representation of the movement quality. Second, to identify the influence of motor and visual expertise on the mental representation of the movement quality of complex gymnastics skills. This investigation was performed using a cluster analysis and its resulting tree diagram as well as a correlation analysis. If the judgment of the movement quality of gymnastics skills is based on the holistic kinematics pattern of the skills, the tree diagram of the pattern of the perceived movement quality and the tree diagram of the kinematics pattern of different performances of the skill should show a similar pattern. If the pattern of the perceived movement quality is similar between gymnastics motor experts, gymnastics visual experts, and novices, the resulting tree diagrams should show a similar pattern.

Methods

Participants

In total, $N = 30$ participants with different levels of gymnastics expertise were recruited for the study. The number of participants was derived from a power analysis when expecting a medium effect (Cohen's f > 0.25, type I error probability 5 %, type II error probability 20%) and is comparable to similar studies conducted by Lander and Lange (1992, 1996) as well as Schack and Mechsner (2006). The motor experts group consisted of $n = 10$ participants who had been active in gymnastics as athletes for a minimum of three years and were still active at the time of investigation (age: $M = 18.80$, $SD = 7.28$). The visual experts group included $n = 10$ participants who had been active in gymnastics as judges, or coaches for a minimum of three years and were still active at the time of investigation (age: $M = 46.90$, $SD = 17.57$). The novices group included $n = 10$ participants who were neither active nor had they ever been active in gymnastics, either as athletes, judges or coaches (age: $M = 30.50$, $SD = 14.42$). The study was conducted in compliance with the Helsinki Declaration for Human Research and the International Principles governing research on humans, as well as in compliance with the ethical guidelines of the local ethics committee.

FIGURE 7.1: Illustration of the gymnastics skill handstand - back handspring on the balance beam.

Instruments and Material

Video Recordings

The video sequences from which the kinematic data was extracted and the stimuli for the computer-based paired comparison perception task were generated showing different handstand - back handspring performances, on the balance beam, (Figure 7.1) by ten female, near-expert gymnasts. The video recordings were taken with a digital camera (240 Hz, 1920 × 1080 pixel) at a distance of 15 m from the beam, from the judges' perspective (viewed on a sagittal plane). The video sequences were cut to the desired duration of the movement from the beginning of the initial movement to a stable landing position. For the experiment, eleven handstand - back handspring trials out of a sample of 60 were chosen. The eleven trials were different in their kinematic pattern as well as in terms of movement quality.

Movement Pattern Analysis

To determine the kinematic movement pattern, the horizontal and vertical coordinates of ten points (body landmarks) were digitized for each frame of the video sequences, using the movement analysis software Simi Motion®. A digital filter built into the software was applied for data smoothing. Each of the ten body landmarks was represented by a two-dimensional time series $[x_j(t);\ y_j(t)]$ with $j = 1, 2, 3, ..., j$ (t = time, j = frame number). The time series of the body landmarks were time-normalized and rescaled to the interval [0; 1000] for each trial, and the kinematic angular data was calculated from the position data of the body landmarks.

The kinematics pattern was calculated using a cluster analysis (Mack et al., 2018). For each time course of the joint angles, as well as for the trunk orientation angle between each pair of two handstand - back handspring trials, Euclidean distances were calculated and added together to give one Euclidean distance value for each pair of two handstand - back handspring trials. A value of zero indicated an identical course, whereas the larger the resulting value, the more dissimilar two trials were. The resulting Euclidean distance matrix was evaluated by a hierarchical cluster analysis, using Ward's hierarchical clustering method (Ward, 1963).

Evaluation of Mental Representation

The perceived movement quality was analyzed using a paired comparison test. The results were structured with a cluster analysis and visualized as a tree diagram, in order to compare the resulting tree diagram of the kinematic pattern and the perceived movement quality of the three different experimental groups. Whenever a video was judged to be of better quality than its counterpart, it was scored with one point. If both videos were judged to be of equal quality, neither was scored with a point. This procedure was performed for all results of the 66 paired comparisons from the 30 participants. The Euclidian distances were averaged over the participants of each experimental group, resulting in three Euclidian distance matrices and three tree diagrams for each of the experimental groups. The approach is similar to the SDA method (Lander, 1991; Lander & Lange, 1996), the SDA-M method (Schack, 2004) or the Repertory Grid Technique (Fransella, Bell, & Bannister, 2004), which are used to identify the way people construe a particular experience.

Procedure

All participants were asked for their written agreement and level of experience before their participation, which was voluntary. Prior to the experiment, each participant received an introduction to the study procedure. In the first step, the participants performed the paired comparison test. Thereby, each video sequence was compared with each of the other video sequences, but not with itself, yielding 66 comparisons. The participants were instructed to decide whether the movement quality was better for the first video sequence presented or the second video sequence presented, or whether both video sequences showed the same movement quality for each comparison. The order of the trials was randomized for each participant within the experiment.

Data Analysis

The Euclidean distance matrices of the kinematic pattern and the perceived movement quality of the three experimental groups were normalized and compared in R, using the dendextend R package (Galili, 2015). Therefore, Baker's Gamma Index, which is a measure of the association between two dendrograms of a cluster analysis, was calculated for all six combinations (Baker, 1994). Therefore, both the merging and the height of the branch on which they merge, was taken into account. The value ranges between -1 and 1. A value near zero implies that the two tree diagrams are entirely dissimilar.

Results

Figure 7.2 shows the tree diagram of the kinematic pattern of the eleven performances of the gymnastic skill handstand - back handspring as well as the resulting pattern of the paired comparison test concerning the perceived movement quality of the three experimental groups. An analysis of the data concerning the number of subgroups (i.e. clusters), with the elbow method and silhouette method, revealed two clusters for the dendrogram of the visual and motor experts, and three clusters for the dendrogram of the novices and the kinematic pattern. Differences between the diagrams of the three experimental groups and the kinematic pattern could also be found in terms of both the height of the branches in which they merged and the allocation of the trials to the two different clusters. The Baker's Gamma Index revealed no evidence that the dendrogram of the kinematic pattern is similar to the dendrogram of the perceived movement quality of the visual experts (r = .353, p = .04), for either the motor experts (r = .142, p = .15) or the novices (r = .154, p = .16). Concerning the comparison of the dendrograms of the perceived movement quality of the different groups, there was a significant correlation between the dendrograms of the visual experts and the motor experts (r = .622, p = .02), and there was also a significant correlation between the dendrograms of the visual experts and the novices (r = .678, p = .01). There was no significant correlation between the dendrograms of the motor experts and the novices (r = .315, p = .178).

Discussion

The study aimed to investigate both the relationship between the perceived movement quality of a gymnastics skill and its kinematics pattern, and the influence of expertise on this relationship. The study built on previous research that has found links between the mental representation of a movement and the movement kinematics (Schack, 2003), together with the results of studies concerning the perceptual significance of certain structural and kinematic aspects in movements when identifying and evaluating human motion (e.g. Troje, 2002). While there is a vast expanse of research investigating the order and structure of functional and sensory features of mental representations, the question remains as to whether all those features were taken into account equally when evaluating the quality of a particular movement, and what role expertise played in those perceptual and cognitive processes. The methodology used here was a first

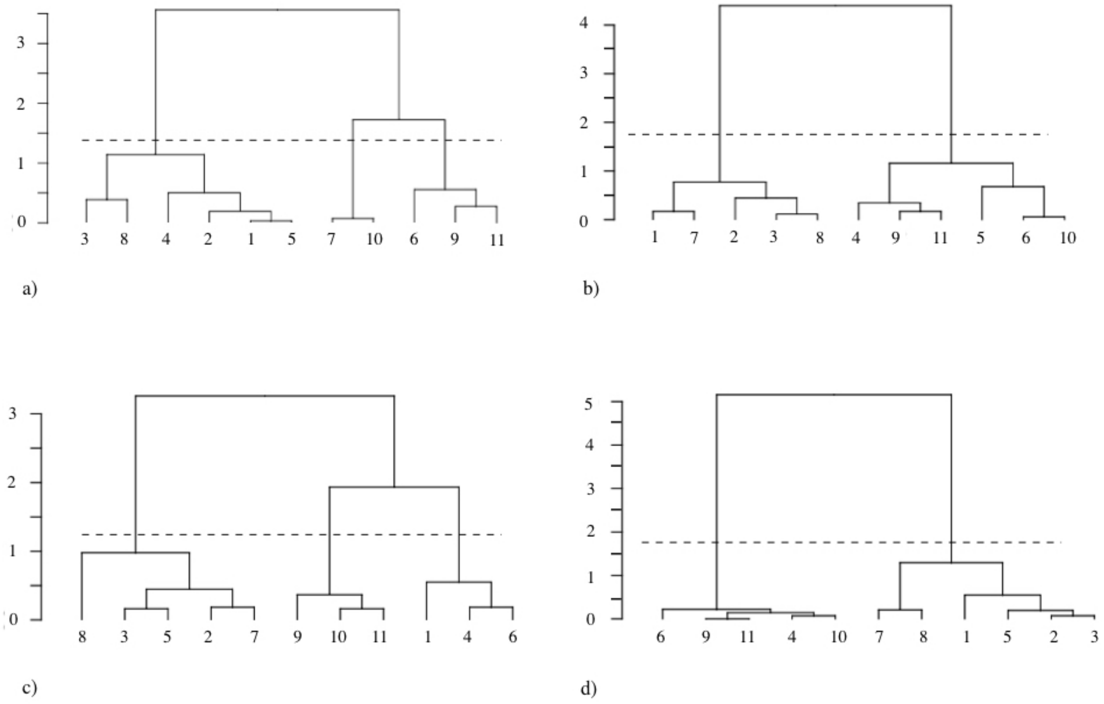

FIGURE 7.2: The tree diagrams resulting from a cluster analysis using Ward's clustering algorithm for a) the kinematic pattern of the respective gymnastics skills as well as the resulting pattern of the paired comparison test concerning the perceived movement quality of b) the motor experts, c) the novices and d) the visual experts. *Note*: The y-axes of the dendrograms represent the measure of closeness (similarity) of individual trials and clusters. The dashed line represents the Euclidean distance below which the clusters are identified.

approach to comparing the structure of the kinematic pattern of a gymnastics skill in its entirety (over time and space), with the mental representation of its qualitative evaluation by participants with different levels of gymnastics expertise. By means of the hierarchical cluster analysis, the resulting structure of the perceived movement quality of each experimental group, as well as the kinematics pattern, was displayed and visualized, inspired by approaches such as the SDA method (Lander & Lange, 1996), the SDA-M method (Schack, 2004) or the Repertory Grid Technique (Fransella et al., 2004). For the tree diagram of the kinematics pattern, it could be assumed that trials that were visualized close to each other have a more similar kinematic pattern than trials that were not as close together. For the tree diagram of the perceived movement quality, it could be assumed that trials that were visualized close together were perceived as being more similar in their movement quality than trials that were not as close to each other.

A comparison of the relationship between the cluster solutions of the three experimental groups with the kinematic pattern revealed no significant correlation. Therefore, it could be assumed that the pattern of the perceived movement quality of the skills was not related to their holistic kinematic pattern. It could be presumed that the quality evaluation is not solely based on kinematic features, but also on other variables. Several studies showed that judges not only base their rating on the basis of kinematics features, but also consider subtle features such as dress, facial expression, mental capacities, personal experiences, expertise, and the point of view and the order of presentation of the gymnasts (Ansorge et al. 1978; de Bruine, 2006; Heinen et al., 2013; Pizzera, 2012; Pizzera & Raab, 2012; Plessner & Haar, 2006; Plessner & Schallis, 2005; Ste-Marie, 1999).

Research showed, that motion recognition is not a matter of single features but rather a process with a holistic character (Troje, 2002). Thereby, there is still a lack of research investigating the different weighting of particular aspects in the process of quality evaluation. A bent knee angle could, for example, decrease the judgment score much more than a bent hip angle, or the landing position could influence the judgment score more than the starting position. To gain insight into those variables that affect quality judgment, analyzing methodologies such as neural network algorithms could be taken into account.

A comparison of the relationship between the tree diagrams of the three experimental groups revealed a significant correlation between the dendrograms of the visual experts and the motor experts, as well as between the dendrograms of the visual experts and the novices, but no significant correlation between the dendrograms of the motor experts and novices. It could be assumed that those differences emerged because the different gymnastics expertise level leads to differences concerning the importance of particular kinematic aspects. Perhaps, for motor experts, the kinematics aspects, which are relevant for a movement to be performed in a certain way, are of central importance, and for novices, they represent the clearly visible quality-determining kinematics aspects. Visual experts, such as coaches and judges, have to keep both the movement-guiding kinematics aspects and the quality-determining kinematic aspects in mind. These results are in line with research on visual search strategies while evaluating gymnastics movements. It could be shown that different levels of gymnastics expertise lead to a different perpetual pattern (focusing on various information sources) as well as different judgments (Bard et al., 1980; Del Campo & Gracia, 2018; Pizzera et al., 2018).

When looking at the number of clusters in the dendrograms of the three experimental groups, two clusters could be identified for both the motor experts and the visual experts, and three clusters could be recognized for the novices. However, not only were differences found concerning the number of clusters, but also for the closeness of the different clusters. The two distinct clusters of the dendrograms of the visual and motor experts were less close than the three different clusters of the dendrograms of the novices. This finding suggests that there is more variance in the decision regarding the movement quality for the novices than for the visual and motor experts. This difference could occur for two reasons. First, it could be that the representation of the kinematic movement pattern of novices was not as detailed as for the visual and motor experts, or second, they had a detailed representation of the kinematic pattern but the representation of what should be rated as good or bad quality was not that detailed. When looking at the results, it could not be stated whether the perception of the movement or the different weighting of the kinematics aspects led to the differences in the dendrograms of the motor experts and the novices. Therefore, it would be interesting to compare the mental representation of a certain gymnastics skill with the pattern of the perceived movement quality. It could be assumed that the mental representation has a mediating influence on the evaluation of the quality of a movement and the quality-influencing variables.

Conclusion

The study revealed differences between visual experts, motor experts and novices in gymnastics, regarding the pattern of the perceived movement quality of the gymnastics skill handstand - back handspring. The pattern of perceived movement quality was not correlated with the holistic kinematic pattern of the judged skills. The results suggest perceptual and cognitive differences for the participants when evaluating gymnastics skills, based on their different previous visual and motor experiences. Investigating the relationship between the pattern of the perceived movement quality and the kinematic pattern may be useful for diagnosing the objectivity of the evaluation of movement quality or guiding the training of judges. Coaches could set a priority on these aspects in training, and judges could keep their own prioritization in mind to avoid a subjective judgment.

Chapter 8

Publication IV

The study described in this chapter has been published following peer review. Full reference details are: **Mack, M. (2020). Exploring cognitive and perceptual judgment processes in gymnastics by using essential kinematics information. *Advances in Cognitive Psychology, 16*(1), 34-44.** The author reserves the right that parts are slightly modified.

Abstract

The aim of the present study was to transfer the method of using the isolated kinematics information combined with the examination of the perceptual-cognitive processes to gymnastics judging and thereby investigating the informational underpinning of skilled perception and judgment. More specifically, a combination of process-tracing methods that include both the gaze pattern (via eye tracking) and the performance judgment (via ratings on a six-point Likert scale) of participants with different levels of gymnastics expertise (visual experts, $n = 14$, motor experts, $n = 17$, novices, $n = 18$) was employed for gymnastics performances of a floor routine (round off, back handspring, back layout somersault), which were shown as videos in original or in stick-figure format. Gaze pattern was thereby analyzed for the whole floor routine as well as for the particular spatial-temporal windows of the three motor skills. Differences between visual experts, motor experts, and novices could be found concerning the judgment score, $\eta_p^2 = .242$ and the accuracy of judgment accuracy, $\eta_p^2 = .196$ but not concerning the gaze pattern. The significant interaction effects between skill and format for the gaze pattern show the importance of the last skill in the judgment processes. Further research should investigate the influence of the judgment instruction on the gaze pattern as well as the importance of the last skill for the judgment score.

Introduction

The judgment of gymnastics skills, often under conditions of limited time and restricted resources, seems to be a complex task, and the performance varies widely (Mercier & Heiniger, 2018). For a reliable and fair evaluation of the performed skill, it is important to know when to focus on the athlete's body; but the movements are fast and the visual display is both large and crowded with information, which is relevant and also irrelevant to the task (Mann et al., 2007). Because there is still a limited understanding of what perceptual information provides the basis for accurate judgment, the overall goal of the present study was to investigate the informational underpinning of skilled perception and judgment. To better understand these processes, a combination of methods that include both the performer's gaze pattern (via eye tracking) and their performance judgment (via ratings on a six-point Likert scale) was employed for gymnastics performances, which were shown in original format, or in stick-figure format, which provides only the essential kinematics information and no surface information.

Artistic Gymnastics is an aesthetic sport, involving very complex and technically challenging sequences of prescribed skills, but also artistic elements. Like other gymnastics disciplines, synchronized swimming or figure skating, the performance depends exclusively on the judges' evaluation, based on criteria defined in the code of points (Arkaev & Suchilin, 2004; FIG, 2017). This is a regular compendium that provides the kinematics criteria by which the skills should be evaluated. Therefore, the kinematics of a performed skill have to be compared with the kinematics of an optimal execution model (Ste-Marie, 1999). The variability and its influencing variables of gymnastics judgments are described and investigated in a large body of literature, in order to identify and better understand the variables in the evaluation process where bias could occur (Ansorge et al., 1978; Heinen et al., 2013; MacMahon & Plessner, 2008; Pizzera, 2012; Pizzera & Raab, 2012; Plessner & Haar, 2006; Plessner & Schallies, 2005; Ste-Marie, 1999, 2000). The complex process of observing and evaluating gymnastics skills is located in the interaction of perception and action (Dosseville et al., 2014; Prinz, 1997), which raises the key question of how the quantifiable movement execution is related to the evaluation of performance quality. Mixed method approaches that provide the opportunity to link the investigation of those different variables, and thereby enable better understanding of their relationship to one another, are still rare. The term 'mixed methods' thereby refers to the method perspective, where mixed methods are viewed, in order to collect, analyze and interpret multiple types of quantitative and qualitative data (Creswell & Tashakkori, 2007).

In general, it is supposed that in the process of perceiving and evaluating, relative information, imparted through the motion itself, is used. More specifically, kinematic cues, over time and space, were integrated into the evaluation process (Johansson, 1973; Troje, 2002). There is also an expanse of research addressing the relationship between movement kinematics and judges' scores in gymnastics. Some studies determined the kinematics predictors affecting the final score out of a large set of potential predictors by using deterministic models (for example: Farana & Vaverka, 2012; Farana et al., 2015; Takei, 1990; 1992, 1998, 2007; Takei et al., 2000). Thus, it might be assumed that kinematics are of mere importance for the perception and evaluation of gymnastics skills. The judges must be able to efficiently and effectively identify the relevant information, direct their attention towards it, and infer its meaning (Williams et al., 1999). Williams and Ericsson (2005) concluded, that such expert performance is possible through the interplay between certain visual search strategies, the anticipation of extended movement cues, recognition of movement patterns and the use of probabilistic information.

One method of investigating the influence of kinematics in the evaluation process is the use of point-light displays. Thereby, the major joints of a moving person are represented by dots. Research has shown that not only particular actions (Dittrich, 1993), but also the gender of a walking person (Barclay, Cutting, & Kozlowski, 1978; Cutting, 1978; Troje, 2002) can be identified through a point-light display, as well as it being possible to identify individual persons (Cutting & Kozlowski, 1977; Stevenage et al., 1999; Troje et al., 2005). The method was also used in the sporting context, where dynamic information was manipulated or occluded and presented to the participants. It could be shown that for the evaluation, relevant information could be extracted from the point-light displays (Cañal-Bruland & Williams, 2010; Huys et al., 2008; Williams, Huys, Cañal-Bruland, & Hagemann, 2009). Thereby, the importance of picking up postural cues and biological motion when attempting to anticipate the opponents next action is highlighted.

The investigation of gaze pattern seems to be an appropriate methodology to better understand visual attention and the underlying cognitive processes in an information search (Vickers, 2007). It can provide important insights into the information used and the importance of various variables in judgment situations, such as in gymnastics judging (Bard et al., 1980; Del Campo & Gracia, 2017; Pizzera et al., 2018). Sport performers must be able to identify the areas with the relevant information, direct their attention, extract this information efficiently and effectively, and interpret its meaning (Williams et al., 1999).

Numerous researchers have used eye tracking systems to record gaze patterns as athletes attempt to anticipate or judge skill performance in both laboratory and field settings (Mann et al., 2019). Therefore, several indices of gaze could be found, which seem to influence performance by indexing an individual's point of interest and information processing, for example the location, and the number and duration of fixations. Those indices were used to operationalize certain information-processing assumptions. For example, a longer duration of fixation seems to lead to more detailed information processing (but not necessarily of the fixated visual cues), by allowing the extraction of an larger volume of information from the visual display (Mann et al., 2007). The location, the point of time and the amount of fixations could give some indication of selective attention allocation (Gegenfurtner et al., 2011).

In an abundance of studies, it could be shown, that there are differences in terms of expertise concerning the judgment of performances in gymnastics, but also in the encoding, processing and retrieving of information (Hars & Calmels, 2007; Heinen et al., 2012; Ste-Marie et al., 2001). Expert judges are better at error detection (Bard, et al., 1980) and perception of body angles (Plessner & Schallies, 2005) compared to novices. But also the perception-action link (common coding theory, Prinz, 1990), that not only visual but also bodily experiences influence the way of perceiving and judging, should be highlighted and taken into account (Pizzera & Raab, 2012). For instance, it could be shown that the motoric knowledge of a skill leads to a more accurate estimation of joint angles and detection of form errors than mere visual knowledge (Pizzera, 2012).

In terms of the underlying mechanisms of those differences, it is proposed that experts are better able to extract information distributed globally over the entire body then novices can (Abernethy, Zawi, & Jackson, 2008). Evidence for a general strong link between perceptual expertise and the kinematics of the action that is viewed is derived from studies demonstrating expert-novice differences in the time-course of information-pick up (Abernethy, Gill, Parks, & Packer, 2001) and anticipatory performances (Ward et al., 2002, Abernethy & Zawi, 2007), by using point-light displays.

The investigation of gaze pattern showed that different levels of gymnastics expertise lead to a different pattern as well as to different judgments. The participants focused on different information sources when making judgments on gymnastics performance, based on their different previous visual and motor experiences. (Bard et al., 1980; Del Campo & Gracia, 2017; Moreno et al., 2002; Pizzera et al., 2018). In a meta-analysis of the underlying mechanism of perceptual skills in sport, Mann et al. (2007) could show that, in general, experts are better at detecting relevant cues by using fewer fixations of longer durations. The results were moderated by variables such as sport type, research paradigm, and presented stimuli. In general, it could be said, that skill-based differences in gaze patterns could be found with experts scanning the display more systematically and fixating on different areas on the display, compared to novices. But the investigation of gaze pattern poses limitations, because of peripherical perception. The point upon which the eyes are fixated is not necessarily the point of attention. This leads to the need of the simultaneous investigation of other variables, such as the judgment results.

To summarize, the complex process of observing and evaluating gymnastics skills is located in the interaction of perception and action (Dosseville et al., 2014; Prinz, 1997). Thereby, the question arises, as to how the quantifiable execution of movement is related to the evaluation of performance quality.

At the moment, there is a growing interest in mixed methods in the field of sport science, that provide the opportunity to link the investigation of those different variables, thereby enabling a better understanding of their relationship (Poizat, Bourbousson, Saury, & Sève, 2012). Visual search patterns and performance judgment were both expertise related variables. But the link between them is still unclear. Because the kinematic pattern should contain the aspects determining performance, the influence of surface information should also be considered. Therefore, the purpose of the current study was to transfer the method of using the isolated kinematics information combined with the examination of the perceptual-cognitive processes to gymnastics judging in a mixed methods approach and thereby describing an approach investigating the informational underpinnings of skilled perception and judgment. More specifically, a combination of methods that include both the gaze patterns (via eye tracking) and the performance judgments (via ratings on a six-point Likert scale) by participants with different gymnastics expertise was employed for gymnastics performances of a floor routine which were shown in original or in stick-figure format, This is a new approach in the field of gymnastics judgment.

Method

Participants

In total, N = 49 participants with different levels of gymnastics expertise were recruited for the study. The number of participants was derived from a power analysis when expecting a medium effect for the main effect of the factor experimental groups (Cohen's f > 0.25, type I error probability 5%, type II error probability 20%). They were required to have normal or corrected-to-normal vision. Dependent on their gymnastics expertise, they were allocated to three different groups (motor experts, visual experts and novices). This was done to control the influence of both visual and bodily experiences on the method of perception and judgment (Pizzera & Raab, 2012). The motor experts group ($n = 17$) included persons (female: $n = 10$; age: $M = 25.71$, $SD = 13.70$, range = 12-56), who had been active in gymnastics as athletes fora minimum of three years and were still active at that date. The visual experts group (n = 14) included persons (female: $n = 9$; age: $M = 31.86$, $SD = 14.26$, range = 19-75), who had been active in gymnastics as judges or trainers for a minimum of three years and were still active at that date. The novices group ($n = 18$) included persons (female: n = 9; age: $M = 30.89$, $SD = 7.43$, range = 12-44), who were not active in gymnastics and had never been active before, either as athletes or as judges or trainers. The age, $F(2,46) = 1.166$, $p = .321$, and gender, $H(1) = 0.017$, $p = .9$, did not differ significantly between the different groups. The participants took part in this study voluntarily and provided informed (parental) consent prior to the study. The study was conducted in line with the ethical guidelines of the local ethics committee, and in compliance with the Declaration of Helsinki for human research and the international principles governing research on humans.

Materials and Instruments

Video stimuli

Participants were asked to judge 40 videos of gymnastics floor routines (round off, back handspring, back-layout somersault). The videos consisted of 20 different trials, which were presented in original and stick-figure format. The floor routines were performed by ten high-level female gymnasts ($M_{\text{age}} = 11.50 \pm 1.43$ years) as they would do in a competition. This was done in order to ensure different performances and different surface information in the video stimuli. The camera was placed approximately 15 meters away from the place where the skills were performed, with its optical axis orthogonally to the direction of motion of the gymnast, simulating the judge's perspective. The stick figure videos were created by using the software Simi Motion. Therefore, a two-dimensional body model, with the x and the y coordinates of the following body landmarks, was created: the forward section of the foot, ankle, knee, hip, shoulder, elbow, wrist, hand (in each case, right and left sides of the body) and head. The videos start with the gymnasts' last step of the take-off phase of the round off and end shortly before the final landing position of the back layout somersault, in order to avoid the influence of the final routine position on the judgment. At the beginning of each video, a short video sequence was inserted, showing the numbers "4", "3", "2", "1", "+", in sequence, at intervals of one second, on the left side of the screen where the floor routine starts. This was done to focus the participants' attention on the place and time at which the first skill starts (Figure 8.1).

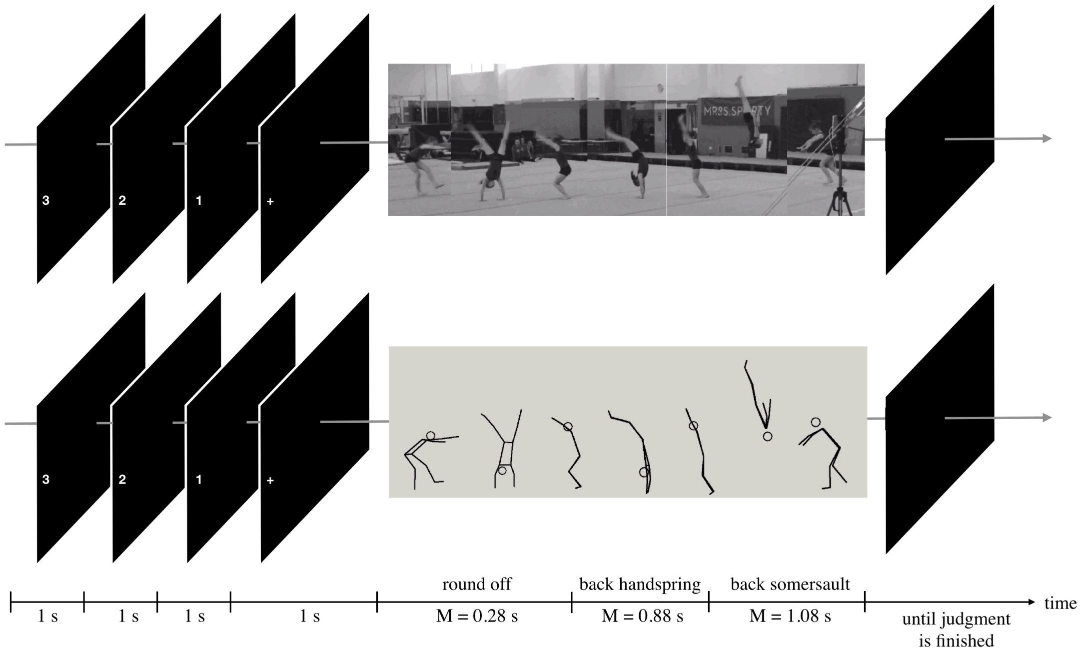

FIGURE 8.1: Example of the sequence of an experimental trial in the original format and in the stick-figure format.

Judgment Performance

Participants were asked to rate each of the videos on a visual analog scale that was anchored to six points, according to the judgment guidelines of the German Gymnastics Federation for young gymnastics talents (DTB, 2001): (1) technically poor execution or fall, (2) technical errors with major postural errors, (3) technical errors with minor postural errors, (4) good execution with minor postural errors, (5) good execution, (6) very good execution. The scale was presented on a tablet, which could be placed in a self-determined comfortable position in front of the screen. On the scale, the six points were marked, and several intermediate stages ($n = 10$) could also be chosen. Each video was shown only once.

To calculate the performance accuracy, the absolute difference of the judgment score to the true scores were calculated. To determine the true sores, the same videos were rated on the same scale by three gymnastics experts (visual and motoric experience in research and practice for over ten years). Thereby, they were allowed to regard the videos at their preferred pace, as often as they wanted to, and to stop them whenever they chose. To evaluate the reliability of the true scores, an intraclass correlation was calculated over all trials ($ICC = 0.7$) as well as separately for the trials in the original format ($ICC = 0.729$) and the stick-figure format ($ICC = 0.664$). These values are all indicative of moderate (Koo & Li, 2016) to good (Cicchetti, 1994) reliability.

Gaze Parameters

The Gazepoint GP3 eye tracker was used to collect gaze data at 150 Hz. The videos were displayed on a 25 inch monitor with a display resolution of 1920 × 1080 pixels. Participants were seated approximately 65 cm from the display. Eye-tracking data was collected from the start of the trials until the end. Three components of gaze behavior were recorded, namely, the average fixation duration, the number of fixations, and the summarized fixation duration. The average fixation duration referred to the average length of time for which a participant fixated on a certain location on the screen, starting at a minimum of 50 ms (Galley & Biniossek, 2015) The number of fixations referred to the number of times a participant fixated on a point on the screen after moving the eyes. The summarized fixation duration referred to the total length of fixations and was calculated by adding the average fixation duration over the total number of fixations. After removing trials in which less than 60% of the eye tracking data was valid for the sequence of the action, the three components of gaze behavior were calculated for the entire floor routine and for particular temporal-spatial windows of the different skills. These only accounted for the gaze parameters for the time and space in which a particular skill was performed. For the round off, this means that the gaze parameters shown were, on average, in the first 13.09 cm (SD = 1.20 cm) of the screen and the first 0.28 s (SD = 0.01 s) of the duration of action. For the back handspring, this means that the gaze parameters shown were, on average, in the first 7.14 cm (SD = 0.62 cm) of the screen and the first 0.88 s (SD = 0.02 s) of the duration of action. For the somersault, this means the gaze parameters that were shown were, on average, in the first 8.12 cm (SD = 1.08 cm) of the screen and the first 1.08 s (SD = 0.28 s) of the duration of action.

Procedure

The experiment lasted approximately 30 minutes, and each participant was tested individually. At the beginning, the participant was informed about the procedure for the experiment and the judgment criteria. Afterwards, for a detailed description of the different experimental groups, gender, age, and gymnastics expertise-related variables were measured. To adapt the calculation of the individual gaze pattern, a calibration routine was carried out. After four familiarization trials, the participant rated the 40 randomly presented videos while having his/her gaze patterns recorded. After the experiment, the participant was debriefed and dismissed.

Analyses

The free statistics software R was used for further data processing and analysis. The significance criterion of $p < .05$ was established for all analyses. The dependent variables of judgment score and judgment accuracy were normally distributed. The distributions of the dependent variables of average fixation duration, summarized fixation duration, and number of fixations differed from the norm, which is typical for gaze data (Galley, Betz, & Biniossek, 2015). The dependent variables were averaged over the 40 video trials. For the analysis of the judgments, two separate 3 (Group) $\times$ 2 (Format) two-way repeated-measures analyses of variance (ANOVAs) were conducted for the dependent variables of judgment score and judgment accuracy. For the analysis of the gaze parameters of the whole floor routine, three separate 3 (Group) $\times$ 2 (Format) two-way repeated-measures ANOVAs were conducted for the dependent variables of average duration of fixation, summarized duration, and number of fixations. For the analysis of the gaze parameters of the temporal-spatial windows of the different skills, three separate 3 (Group) $\times$ 2 (Format) $\times$ 3 (Skill) three-way repeated-measures ANOVAs were conducted for the dependent variables of average fixation duration, summarized fixation duration, and number of fixations, for each of the three skills. Both the gaze pattern and the judgment measurements (judgment score and accuracy) were assessed and analyzed, because both aspects have their limitations as well as advantages, thereby complementing one another. Concerning the judgment measurements, both the score and the accuracy were measured and analyzed, because both are crucial factors for the reliability and validity of performance judgment. The gaze pattern was analyzed separately with regard to the entire floor routine and the three skills, in order to facilitate insights into how the gaze pattern varies over a gymnastics skill routine and how it could, therefore, influence the weighting of the different skills in the overall judgment score.

Results

Judgement

Table 8.1 shows the descriptive statistics of the judgment score and the judgment accuracy, broken down by experimental groups and formats. Two 3 (Group) $\times$ 2 (Format) two-way repeated-measures ANOVAs were performed for the dependent variables of judgment score and judgment accuracy. Based on the histogram of residuals - both the residuals of the judgment score as well as those of accuracy were assumed as being normally distributed. Levene's test indicated equal variances for the judgment score, $F = 1.23$, $p = .303$, and unequal variances for the judgment accuracy, $F = 2.58$, $p = .031$.

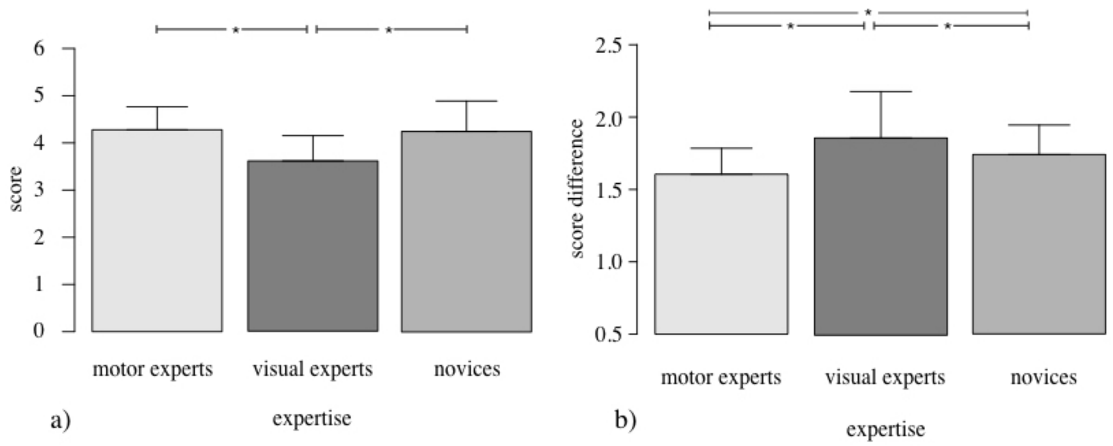

FIGURE 8.2: Means (M) and standard deviations (SD) of a) the judgment score and c) the accuracy in dependence of expertise. *Note.* The * indicate the significant differences.

Judgment score: For judgment score, the ANOVA revealed a significant main effect for expertise, $F(2{,}46) = 7.348$, $p = .002$, $\eta_p^2 = .242$, and format, $F(1{,}46) = 13.269$, p < .001, $\eta_p^2 = .224$, but no significant interaction effect, $F(2{,}46) = 0.706$, $p = .499$, $\eta_p^2 = .03$. Skills which were presented in the stick figure format were rated higher than skills that were presented in the original format. A post hoc t-test with Holm correction showed that both motor experts and novices rated the skills better than visual experts ($p < .001$) (Figure 8.2).

Judgement accuracy: For judgment accuracy, the ANOVA revealed a significant main effect for expertise, $F(2{,}46) = 5.617$, $p = .007$, $\eta_p^2 = .196$, but no significant main effect for format, $F(1{,}46) = 0.005$, $p = .941$, $\eta_p^2 < .001$, and no significant interaction effect, $F(2{,}46) = 2.297$, $p = .112$, $\eta_p^2 = .048$. A post hoc t-test with Holm correction showed a higher judgment accuracy for both motor experts and novices than for visual experts ($p < .01$) and a higher judgment accuracy for novices than for motor experts ($p = .024$) (Figure 8.2)

Gaze Parameters

Whole floor routine

Table 8.1 shows the descriptive statistics of the gaze parameters. To investigate the gaze patterns for the floor routines as a whole, three 3 (Group) × 2 (Format) two-way repeated-measures ANOVAs were performed for the dependent variables of average fixation duration, summarized fixation duration, and number of fixations. Based on the histogram of residuals, both the residuals of the average fixation duration as well as of the number of fixations can be assumed as not being normally distributed, which is typical for gaze data (Galley et al., 2015). The residuals of the summarized fixation can be assumed as being normally distributed, based on the histogram of the residuals. Levene's test indicated equal variances for the average fixation duration, $F = 0.89$, $p = .493$, for the summarized fixation duration, $F = 0.40$, $p = .847$, and for the number of fixations, $F = 0.43$, $p = .823$.

Average fixation duration: For the variable average fixation duration, the ANOVA revealed a significant main effect for format, $F(1,46) = 17.001$, $p < .001$, $\eta_p^2 = .270$. All other effects were not significant ($p > .05$). Skills which were presented in the stick-figure format ($M = 0.484$, $SD = 0.154$) lead to a higher average fixation duration than skills that were presented in the original format ($M = 0.432$, $SD = 0.105$).

Summarized fixation duration: No significant effects could be revealed by the ANOVA for the variable summarized fixation duration (p > .05).

Number of fixations: For the number of fixations, the ANOVA revealed a significant main effect for format, F (1,46) = 17.505, p < .001, $\eta_p^2 = .276$, and no other significant effects (p > .05). Skills which were presented in the original format ($M = 6.296$, $SD = 0.986$) led to a higher number of fixations than skills that were presented in the stick-figure format ($M = 5.850$, $SD = 1.122$).

Temporal-spatial windows of the skills

To investigate the gaze pattern for the floor routines, separated into the three different skills, three 3 (Group) × 2 (Format) × 3 (Skill) three-way repeated-measures ANOVAs were conducted for the dependent variables of average fixation duration, summarized fixation duration, and number of fixations.

Average fixation duration: For the average fixation duration, the ANOVA revealed a significant main effect of skill, $F(2,92) = 32.061$, $p < .001$, $\eta_p^2 = .411$. A post hoc t-test with Holm correction showed a significant difference concerning the summarized fixation duration for all of the three skills ($p < .05$) Thereby the back handspring ($M = 0.154$, $SD = 0.335$) leads to the lowest average fixation duration, followed by the back layout somersault ($M = 0.257$, $SD = 0.154$) and the round off ($M = 0.601$, $SD = 0.422$). None of the other effects were significant ($p > .05$)(Figure 8.1).

Summarized fixation duration: For the summarized fixation duration, the ANOVA revealed a significant main effect of skill, $F(2,92) = 29.002$, $p < .001$, $\eta_p^2 = .387$, a significant main effect for format, $F(1,46) = 11.669$, $p = .001$, $\eta_p^2 = .202$, and a significant interaction effect between skill and format, $F(2,92) = 6.092$, $p = .003$, $\eta_p^2 = .117$ (Figure 8.1). None of the other effects were significant ($p > .05$). A post hoc t-test with Holm correction showed a significant difference concerning the summarized fixation duration for all of the three skills ($p < .05$). Thereby the back handspring ($M = 0.188$, $SD =$

0.387) leads to the lowest summarized fixation duration, followed by the back layout somersault (M = 0.499, SD = 0.236) and the round off (M = 0.660, SD = 0.407). Skills which were presented in the stick-figure format (M = 0.493, SD = 0.411) lead to a higher summarized fixation duration than skills that were presented in the original format (M = 0.405, SD = 0.389). The significant interaction between skill and format indicated that the effect of format on the summarized fixation duration was different for the different skills. A post-hoc t-test revealed that there was a significant influence of format for the back layout somersault (Z = 22, $p < .001$) but not for the round off (Z = 635, p = .633), and the back handspring (Z = 51, p = .139). The summarized fixation duration was significantly higher for the stick-figure format (M = 2.537, SD = 0.578) than for the original format (M = 1.650, SD = 0.453), for the back layout somersault (Figure 8.1)

Number of fixations: For the number of fixations, the ANOVA revealed both a significant main effect of skill, $F(2,92) = 179.194$, $p< .001$, $\eta_p^2 = .796$, and format, $F(1,46) = 53.685$, p $< .001$, $\eta_p^2 = .539$, and a significant interaction effect between skill and format, $F(2,92)$ $= 30.263$, $p< .001$, $\eta_p^2 = .397$ (Figure 8.1). None of the other effects were significant ($p >$ .05). A post hoc t-test with Holm correction showed a significant difference concerning the number of fixations for all of the three skills ($p < .05$). Thereby the back handspring (M = 0.422, SD = 0.724) lead to the lowest number of fixations, followed by the round off (M = 1.193, SD = 0.271) and the back layout somersault (M = 2.093, SD = 0.684). Skills which were presented in the stick-figure format (M = 1.426, SD = 1.028) lead to a higher number of fixations than skills which were presented in the original format (M = 1.047, SD = 0.718). The significant interaction between skill and format indicated that the effect of format on the variable number of fixations was different for the different skills. A post-hoc t-test revealed that there was a significant influence of format for the back handspring (Z = 19.5, p = .007) and the back layout somersault (Z = 15.5, p = $< .001$), but not for the round off (Z = 343, p = .904). The number of fixations was significantly higher for the stick-figure format than for the original format for both the back handspring (stick figure: M = 0.201, SD = 0.320; original: M = 0.176, SD = 0.448) and the back layout somersault (stick-figure: M = 0.616, SD = 0.273; original: M = 0.382, SD = 0.101) (Figure 8.1).

Discussion

The overall goal of the present study was to investigate the informational underpinnings of skilled perception and judgment. To better understand these processes, a combination of methods, including both the participants' gaze patterns (via eye tracking) and

gaze parameters	average fixation duration				summarized fixation duration				number of fixations			
video format	original		stick-figure		original		stick-figure		original		stick-figure	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
overall												
Motor experts	0.405	0.080	0.457	0.102	2.360	0.250	2.362	0.271	6.491	1.060	5.844	0.935
Visual experts	0.457	0.121	0.506	0.191	2.486	0.387	2.451	0.334	6.076	1.034	5.882	1.166
Novices	0.436	0.109	0.487	0.169	2.478	0.350	2.425	0.444	6.282	0.892	5.816	1.300
differentiated by skill												
round off												
motor experts	0.580	0.330	0.510	0.261	0.636	0.306	0.565	0.263	1.202	0.257	1.173	0.177
visual experts	0.583	0.454	0.718	0.661	0.631	0.434	0.782	0.643	1.141	0.460	1.227	0.212
novices	0.622	0.386	0.611	0.434	0.700	0.361	0.658	0.417	1.224	0.235	1.187	0.263
back handspring												
motor experts	0.193	0.420	0.217	0.256	0.234	0.527	0.255	0.312	0.350	0.566	0.572	0.649
visual experts	0.107	0.275	0.159	0.283	0.114	0.291	0.205	0.309	0.236	0.470	0.798	1.082
novices	0.152	0.479	0.091	0.226	0.169	0.483	0.145	0.342	0.296	0.604	0.330	0.818
back somersault												
motor experts	0.235	0.043	0.312	0.227	0.407	0.112	0.652	0.287	1.765	0.504	2.455	0.545
visual experts	0.242	0.038	0.309	0.308	0.381	0.092	0.666	0.383	1.608	0.400	2.564	0.554
novices	0.234	0.042	0.217	0.029	0.360	0.096	0.543	0.110	1.574	0.444	2.592	0.647

TABLE 8.1: Means (M) and standard deviations (SD) of gaze parameters (mean fixation duration, summarized fixation duration, number of fixations) as a function of format (original video, stick-figure video) and expertise (visual experts, motor experts, novices) for the temporal-spatial windows of the floor-routine.

their performance judgments (via ratings on a six-point Likert scale) was employed for gymnastics performances, which were shown as videos in original or stick-figure format. To our knowledge, the results of the current study are the first judgment and gaze pattern results concerning the influence of essential kinematics information and the differing importance of the particular skills in a sequence.

At the moment, there is a growing interest in mixed methods in the field of sport science, in order to gain insight into the topic from different perspectives (Pizzera et al., 2012). The current study contributes with its approach, by evaluating the usefulness of a joint analysis of (a) first-person judgment results, (b) objective eye tracking results and of (c) manipulation of the proportion of surface information.

The analysis of the judgment scores showed that visual experts rated both the videos presented in stick-figure format and in the original format lower than did the motor experts and the novices. This is not surprising when considering the higher level of experience of visual experts, such as judges and coaches, in evaluating and judging skills (Bard, et al., 1980; Plessner & Schallies, 2005). It might be the case that they recognize more and smaller mistakes than motor experts and novices, which leads to lower judgments. The high level of the gymnastics performances shown in the experiment and the associated small errors in the movement executions strengthen this assumption. The abundance of minor movement errors in videos presented in the stick-figure format could also account for the main effect of format, where videos in stick-figure format were rated higher than were the videos in the original format.

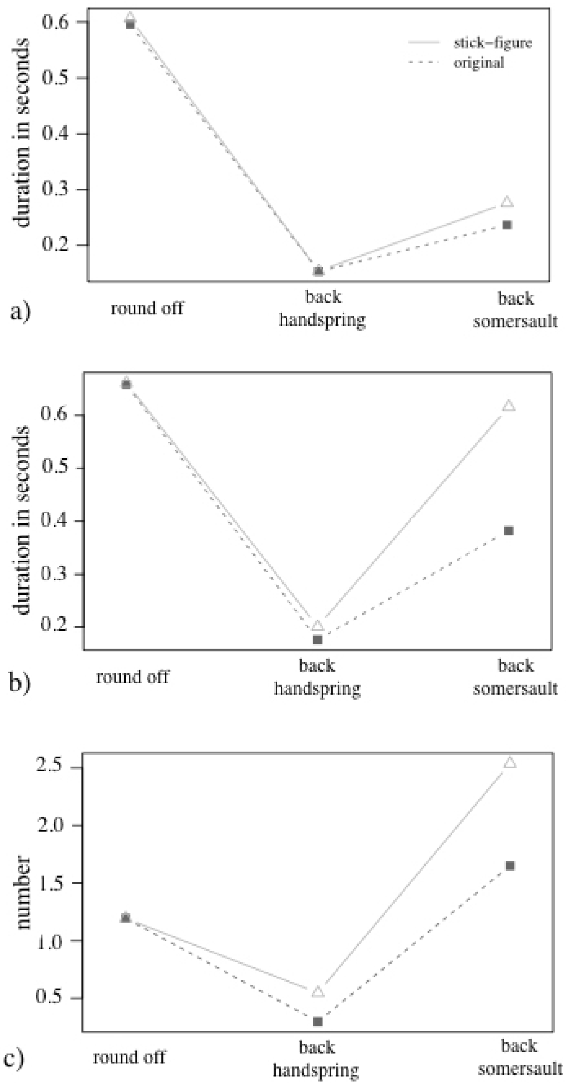

FIGURE 8.3: Interaction plots of the temporal-spatial window of the skills (round off, back handspring, backward somersault) and the format (original, stick-figure) for a) the average fixation duration, b) the summarized fixation duration and c) the number of fixations.

The judgment of the visual experts was not only the lowest, but also the least accurate, followed by the novices and the motor experts, who rated them the most accurately. It is possible that the judgment procedure of coaches and judges usually differs from the task in the current study, and that this caused problems in habituating to the judgment instructions given for the current study. Another explanation is the inclusion of both coaches and judges in the experimental group of visual experts. Both deal with the evaluation of motor skills, but each with a different focus. The task of the judges is to recognize movement errors according to criteria defined in the code of points (Arkaev & Suchilin, 2004; FIG, 2017), whereas coaches have to identify the causes of these errors. They have to relate the errors to the dependent movement characteristics, in order to decide which movements should be modified to remove the execution errors and to gain a higher execution score. Also, the enactive assumption has to be considered in this context (e.g., Hauw et al., 2008). In this theoretical perspective, judgment and decision-making are embodied, so that motor experts may have developed a more precise judgment concerning the kinematics of gymnastics performances. It would be pertinent to consider this account in further research. However, different learning strategies in order to achieve knowledge should also be considered, as they could have led to different processing of movement information (Kermarrec et al., 2004).

Regarding the gaze parameters of the whole floor routine, a main effect of format was found for the average fixation duration and the number of fixations. Performances which were presented in the stick-figure format led to a higher average fixation duration and to a lower number of fixations than did performances presented in the original format. It is assumed that a longer fixation duration leads to more detailed information processing, by allowing for the extraction of a higher amount of information (Mann et al., 2007; Williams et al., 1999). Therefore, the longer average fixation duration for the stick-figure videos could be explained by the lower amount of information shown on the screen. This might have allowed the participants to focus only on the motion patterns, without being distracted by other visual information. This corresponds to the concomitant lower search rate, which is shown in the lower amount of fixations for the stick-figure format in comparison to the original format.

To investigate the gaze pattern in relation to the different skills, we analyzed the gaze patterns shown at the particular temporal-spatial windows of the three different skills. This means that, for instance, for the round off, only the gaze pattern that was shown at the spatial window of the screen where the round off was performed, and only the particular time window in which the round off was performed, were extracted. As a result, the significant main effect of the average fixation duration disappeared, and the significant main effect of the number of fixations became inverted. Thus, performances which were presented in the stick-figure format led to the same average fixation duration, but to a higher number of fixations than did videos presented in the original format. Differences between the global gaze patterns and the patterns of the temporal-spatial window of the particular skills might lead to the assumption that in the stick-figure format, the participants could focus more on the temporal-spatial window of the shown skill and were not occupied with extracting and processing the information from another temporal-spatial window.

Significant interaction effects between skill and format were found for the summarized fixation duration and the number of fixations. Thus, a different gaze pattern between the original format and the stick-figure format could be found, especially for the last skill, the back layout somersault, whereas it was not found for the first skill, the round off. Thus, the stick-figure format led to a higher number of fixations and a longer summarized fixation duration than did the original format. This interaction might lead to the assumption that the last skill plays a particular role in the judgment of the whole floor routine. On the basis of the assumption above that the stick-figure format allows the participants to focus more on the relevant information (Abernethy & Zawi, 2007), it could also be assumed that the last skill is of major importance for the participants in the judgment of an entire floor routine. In further studies, for instance, with the temporal occlusion paradigm as an experimental method, the judgment of the particular skills and their influence on the judgment of an entire floor routine should be investigated. This could lead to more detailed insights as to whether, for instance, the judgment of the whole floor routine is more strongly correlated to the judgment of the last skill than to the first.

Interestingly, against the assumptions that expertise leads to a different gaze pattern (Bard et al., 1980; del Campo & Gracia, 2017; Pizzera et al., 2018), no main effect for expertise could be found. Visual experts, motor experts, and novices showed the same gaze patterns. One explanation for that could, once again, be the given instruction to judge the whole floor routine with one overall judgment score between one and six, which is an unusual judgment procedure for judges and coaches. It would be interesting to investigate, in further studies, the extent to which the gaze pattern is related to the given judgment instruction. Furthermore, in additional studies, the number of participants should be increased to ensure a high power not only for the main effects of the factor (different experimental groups), but also for all the specific simple effects (Brysbaert, 2019).

Conclusion

This experiment used gaze patterns, in combination with isolated kinematics information, as a process tracing measure for judgment in gymnastics, while attempting to maintain a certain level of ecological validity by, for instance, maintaining the viewpoint of the judge's perspective in a competition. Differences between visual experts, motor experts, and novices could be found concerning the judgments, but not in the gaze patterns. Moreover, the results showed the importance of the last skill in the judgment process. Further research should investigate the influence of the judging instructions on the gaze pattern as well as the importance of the last skill for the judgment score.

Chapter 9

Publication V

The study described in this chapter has been published following peer review. Full reference details are: **Mack, M., Bryan, M., Heyer, G., & Heinen, T. (2019). Modeling judges' scores in artistic gymnastics. *The Open Sports Sciences Journal, 12*, 1-9.** The author reserves the right that parts are slightly modified.

Abstract

In artistic gymnastics, performance is observed and evaluated by judges, based on criteria defined in the code of points. However, there are manifold influences discussed in literature, that could potentially bias the judges' evaluations in artistic gymnastics. In this context, several authors claim the necessity for alternative approaches to judging gymnastics, utilizing biomechanical methods. The aim of this study was to develop and evaluate a model-based approach for judging gymnastics performance, based on quantitative kinematic data of the performed skills. Four different model variants, based on kinematic similarity, calculated by a multivariate exploratory approach and the Recurrent Neural Network method, were used to evaluate the relationship between the movement kinematics and the judges' scores. The complete dataset consisted of movement kinematic data and judgment scores of a total of $N = 173$ trials of three different skills and routines from women's artistic gymnastics. The results exhibit a significant relationship between the predicted score and the actual score for six of the twelve model calculations. The different model variants yielded a different prediction performance, in general, across all skills and also in terms of the different skills. Specifically, only the Recurrent Neural Network method exhibited significant correlation values between the actual and the predicted scores for all three investigated skills. The results were discussed in terms of the differences in the models as well as the various factors that might play a role in the evaluation process.

Introduction

In competitive sports, judgment of the performance is of vital importance. Sports performance can be either assessed by objective measurements (e.g. time in running, or points in tennis) or by subjective judgments (e.g. points in artistic gymnastics) (Plessner & Haar, 2006). In artistic gymnastics, performance is observed and evaluated by judges, based on criteria defined in the code of points (Arkaev & Suchilin, 2994; FIG, 2017). Thereby, the assumption is that highly skilled judges come to reliable and fair evaluations of observed performances (Pizzera, 2012). However, there are manifold potential influences discussed in literature, that could potentially bias the judges' evaluations (Giblin, Farrow, Reid, Ball, & Abernethy, 2015; Jeraj, 2016; Leskosek et al., 2010; Ste-Marie, 1999). In this context, several authors claim the necessity for alternative approaches to gymnastics judging (Čuk, 2015; Omorczyk, Nosiadek, Ambroy, & Nosiadek, 2015). A particular emphasis is laid on approaches that utilize biomechanical methods, due to their high degree of measurement precision and reliability (Díaz-Pereira, Gomez-Conde, Escalona, & Olivieri, 2014; Heinen et al., 2012). In this context, one should acknowledge that the International Gymnastics Federation is now also increasing its efforts to implement a complex judging support system, based on movement data supports the demand for an objective and potentially unbiased evaluation of gymnastics performance (www.fujitsu.com, 2017). However, before implementing and road-testing a particular system, there has to be some empirical evaluation of such a system. Therefore, the aim of this study was to develop and evaluate a model-based approach for judging gymnastics performance, based on quantitative kinematic data of the performed skills.

In sports for which performance is assessed by objective measurement, such as time in running or points in tennis, it is argued, in literature, that there is no optimal movement pattern that is associated with a high level of performance. Every expert has his/her own optimal movement pattern that meets the constraints in a given situation (Bauer & Schöllhorn, 1997; Hausken, 2017). In contrast, in artistic gymnastics, particular movement characteristics have to be fulfilled. When judging artistic gymnastics performances during a competition, the judgment is made, based on criteria defined in the code of points, in order to make the judgment as objective and comparable as possible. For instance, if a gymnast performs a skill with bent arms or bent knees, he/she may be penalized with a deduction up to 0.50 points. If he/she touches the landing mat with one or two hands during landing, he/she may be penalized with a deduction of 1.00 point. The final execution score is calculated by averaging three out of the five scores, whereby the highest and lowest scores are omitted from this calculation. It is believed that this averaging procedure functions as a control for outliers (Field, Miles, & Field, 2012).

Artistic gymnastics comprises fast and complex skills, and for the average person, it seems almost impossible to ascertain all of the movement-relevant information required for a fair judgment and evaluation of the observed skill at hand, according to the criteria defined in the code of points. Therefore, judges have to acquire both particular knowledge and specific skills through specialized judging courses (Mac Mahon et al., 2015; Pizzera & Raab, 2012). It is a well known fact that judgments of sports performance are influenced by many variables other than the mere performance, and are thereby often biased. Variables that influence the judgment process are social cognition, information processing, perception or memory processes (Plessner & Haar, 2006). For instance, it could be found that the evaluation by the gymnastic judges of the cross on the rings, which is a static strength element, is influenced by their viewing position. This might lead to biased evaluation scores, especially for floor routines where the skills are executed from different directions on the floor. While the viewing position is one source of judgment bias, there are many other sources in the environment, of which people are even less aware (Plessner & Schallis, 2005). Then, when it comes to the encoding of the perceived information, prior knowledge may have an influence. For instance, novices may perceive a gymnastics floor routine as a random pattern of difficult movements, whereas for expert gymnasts, a gymnastics floor routine is the result of particular skills that should be performed in a prescribed way. Not only prior knowledge, but also cues that evolve from the competitive environment itself, are likely to influence performance judgments. For instance, it could be shown that serial position effects are involved when performances are judged in sequence (Ansorge et al., 1978; Bruine de Bruine, 2006). This may lead to the problem that the judgment of the presented performance may simply be influenced by the starting position of the gymnasts in a competition, and thereby affect the outcome of the competition.

In relevant literature, there is a strong argument for biomechanically-driven judging procedures, especially in technical sports, such as gymnastics (Čuk, 2015). For instance, recently, a system capable of measuring horizontal displacement on the trampoline bed, together with the duration of flight was presented (Ferger & Hackbarth, 2017). The duration of flight and horizontal displacement are two parameters that could potentially be useful in judging trampoline performance; however, movement technique consists of a considerable amount of other information that is not captured by these two parameters (i.e. changes in body posture during flight, number of somersaults and twist rotations, etc.). Another quite innovative approach, using an algorithm, was implemented in real-time computer vision software for rhythmic gymnastics (Díaz-Pereira et al., 2014). This software extracted detailed velocity field information from body movements from video shots or live performance video streams of dance movements. By first analyzing the spatio-temporal trajectories and then comparing them with those stored in a database, the authors could reliably classify the recorded movement as well as calculate a judgement score. The method accurately determined scores for different standard gymnastic movements that were comparable to those determined by expert judges. However, this algorithm may work for simple movements, but not for complex movements such as those found in gymnastics, where both static and dynamic skills as well as twists and somersaults are performed in one routine. To capture all those different movement aspects, one needs complete kinematic information, which implies capturing the performed movement with a high degree of accuracy in all three dimensions. Currently, software solutions exist that allow for holistic and instantaneous data collection and analysis of kinematic information (Colyer, Evans, Cosker, & Salo, 2018). There is an expanse of other comprehensive research that addresses the relationships between movement kinematics of gymnastics skill executions and judges‘ scores (Farana et al. 2015; Farana, & Vaverka, 2012; Takei, 2007; Takei et al., 2000; Takei & Dunn, 1996; Takei, Dunn, Blucker, & Nohara, 2000; Takei & Kim 1990). One common result of the aforementioned research was that some kinematic variables correlated well with the judges’ scores. This is especially the case for the vault exercise, which is very fast, technical and includes no artistic evaluation (Takei, 2007). There are several options for investigating kinematics variables, from time-discrete values of selected variables to time and space continuous values of multiple variables. For example, investigating the skill of a handspring during the vault performance could be achieved by simply analyzing the shoulder angle at the moment when the gymnast first touches the vault after the first flight phase, or by analyzing the interplay between the shoulder and the hip angle during the entire movement.

While the aforementioned approaches provide a strong argument for assessing biomechanical information in the evaluation of gymnastics skills, it might be questionable whether using time-discrete measures would be optimal here (Lees, 2002; Stergiou, 2004). Furthermore, research highlights that the perception of motion is better when dynamic motion information is available, compared to structural information alone (Blake & Shiffrar, 2007; Troje, 2002). Thus, a biomechanical approach to gymnastics judging should capture the course of the skills to be judged, with parameters that are perceivable by external observers, such as spectators who might have only sparse knowledge of biomechanics (Federolf, Tecante, & Nigg, 2012; Massidda & Calò, 2012).

In summary, when the evaluation of movement is based on the perception of movement, and the perception of movement is a holistic process that takes the kinematic pattern of the movement over time and space into account, then gymnastics skill performances with a similar kinematic pattern should be similarly compared to gymnastics skill performances with a different kinematic pattern. The aim of this study was to develop and evaluate a model-based approach for judging gymnastics performance, based on quantitative kinematic data of the performed skills. Because the International Gymnastics Federation is intensifying its efforts to implement a computer-based judging support system, based on movement data, there is a demand for investigating this rationale. To identify the similarities and dissimilarities between the kinematic patterns of different movement realizations, two different methods were used, namely, a multivariate exploratory approach and a Recurrent Neural Network (Hochreiter & Schmidhuber, 1997). Four different model variants, based on the two methods, were used to prove a direct and positive relationship between movement kinematics and judges' scores. Based on the previous findings, it was predicted that experts would evaluate gymnastics skills, based on the kinematic patterns of a particular skill. More specifically, it is assumed, that two skill performances that are structurally identical in terms of movement kinematics should be judged with the same score. The more two skill performances differ in movement kinematics, the more different the scoring should be.

Material and Methods

Description of the Dataset and Model Database

The complete dataset consisted of movement kinematic data and judgement scores of a total of $N = 173$ trials of three different skills from women's artistic gymnastics (round off, back handspring, back layout somersault on floor, $n_1 = 58$; back handspring from handstand position on balance beam, $n_2 = 57$; handspring on vault, $n_3 = 57$; *Note*: Unfortunately, one trial had to be excluded for the balance beam and vault, due to problems with data collection). The number of trials was sufficient regarding an a-priori power analysis when expecting a medium effect (Cohens' f = 0.25, type I error probability = .05, type II error probability = .20). The skills were performed by ten high-level female gymnasts (mean age: 11.50 ± 1.43 years). Each gymnast performed six to seven executions of each skill, which is similar to or even below their normal training workload. The gymnasts were active members of a national high-performance training center and they exhibited an average training volume of more than 25 hours per week. They participated in national and international competitions. All gymnasts were informed about the purpose of the study prior to data collection and participated voluntarily. The data collection took place with parental consent, and the study was conducted in line with the ethical guidelines of the local ethics committee, and in compliance with the Declaration of Helsinki for human research and the international principles governing research on humans. The task was for the participants to perform the aforementioned skills on each apparatus, as they would do in a gymnastics competition. Thus, each apparatus was set up in accordance with the competition guidelines of the International Gymnastics Federation for women's artistic gymnastics (FIG, 2017). The gymnasts' performances were recorded with a digital video camera operating at 240 Hz (spatial resolution: 1920 x 1080 pixels). The camera was placed approximately 15 meters away from the apparatus, with its optical axis orthogonally to the direction of movement of the gymnast, thereby simulating the judge's perspective.

Judgment Scores

All 173 trials were presented to five subjects with high visual gymnastics expertise. All subjects were informed about the purpose of the study and gave their consent prior to data collection. They were asked to rate each of the trials on a visual analog scale that was anchored to six points, according to the judgement guidelines of the German Gymnastics Federation for young gymnastics talents (DTB, 2001). To evaluate the judgment scores, the inter-rater reliability was calculated (see Table 9.1). Finally, the judgment scores were averaged, in order to calculate a final judgement score for each trial of the dataset (FIG, 2017).

Kinematic Data

The movement analysis software Simi Motion® was used for digitizing and analyzing the movement kinematics of each of the trials. Thereby, a two-dimensional body model with the x and the y coordinates of the following body landmarks was determined: the forward section of the foot, ankle, knee, hip, shoulder, elbow, wrist, hand (in each case, right and left sides of the body) and head. The time-series of each digitized body landmark was time-normalized and rescaled to a time interval from zero to 1000; this was done to ensure structural comparability between skills of (slightly) different duration. We calculated the angle time series denoting the x and y coordinates for the flexion and extension of both ankle joints, both knee joints, both hip joints, both shoulder joints, both elbow joints and both wrist joints, as well as the orientation angle of the trunk. The free statistic software R (R Core Team, 2017) was used for further data processing and analysis. The neural networks were created using the Keras framework (Chollet et al., 2015) in combination with tensorflow (Martin et al., 2015), to enable GPU training.

Model Assumptions and Calculations

To predict the judgment score from the gymnasts' movement kinematic data, an approach based on the structural similarity of the movement pattern was used. The approach implied a direct and positive relationship between the movement kinematics and the judges' scores. Thus, one of the main assumptions was that two skill performances that are structurally identical in movement kinematics should be judged with the same score. The more two skill performances differ in movement kinematics, the more the scoring should differ (Figure 9.1).

Structural Similarity

Structural similarity was operationalized using the Euclidean distance or a specific pattern in the neural network. Euclidean distances are a mathematical measure, representing the mathematical distance or similarity between two objects. To get the Euclidean distances, cumulated squared differences were calculated for a particular choice of corresponding joint angle time-series between two trials (Schöllhorn et al., 2013). An eight-segment model of the human body with the corresponding knee joints, hip joints, shoulder joints (in each case, right and left sides), as well as the orientation angle of the trunk were used (Behnke, 2001; Jaitner et al., 2001). Therefore, one skill performance contained information from eight variables. The calculated squared differences were added to give a final Euclidean distance value. A smaller Euclidean distance between two skills indicated a higher degree of kinematic similarity between the two skills, whereas a larger Euclidean distance between two skills indicated a lower degree of kinematic similarity of the two skills. The calculation of the Euclidean distances resulted in one distance matrix for each apparatus.

The Recurrent Neural Network is a specific form of an artificial neural network (Marsland, 2015). The network was trained to imitate the judges' scoring. Neural networks are a form of machine learning, of which the architecture of the model is roughly based on the human brain. The model consists of inputs and outputs, which are connected by so-called axons and neurons. Similar to the human nervous system, a neuron fires when a specific amount of energy has reached it, and by firing, a signal is passed to one or more other neurons. In an artificial neural network, neurons are represented by nodes in a layer, and the signals are passed via so-called weights. Information is given to the network in the form of numerical values. These values are passed through the network using the weights and neurons. The output of the network is also in a numerical form, which means that these values have to be interpreted depending on the use case.

For the given use case, the network was given the joint angles of the body as well as the absolute position of the joints in the video frame. The data was entered into the network sequentially. The joint angles were normalized using a sine function. This created two advantages: first, the data was in a fixed range of -1 to 1, and second, the data kept numerically distant values close, such as the angles 359° and 1°. The target values of the judges' scores were normalized to values between 0 and 1, by dividing them by 100. A GRU was used as an RNN layer type (Chung, Gulcehre, Cho, & Bengio, 2014).

Model Variants

On the basis of these approaches, five different model variants were developed, in order to simulate the judging process. With the five models, the full bandwidth of possibilities of judging was covered, from taking the order of the judgments to taking five judgments and removing the best and the worst, to taking only the kinematic pattern. This resulted in the following five model variants: 1) Model variant #1 *best/worst*: We selected the two trials from the database, with the best and the worst final judgment score. The predicted score of a particular evaluation trial was calculated as the weighted average of the best and worst scores, thereby using the Euclidean distances between the evaluation trial and the best/worst trials as the weights for the calculation. 2) Model variant #2 *nearest neighbor*: The database trial with the lowest Euclidean distance to the evaluation trial (i.e. the greatest similarity between both trials in terms of movement kinematics) defined the score of the evaluation. 3) Model variant #3 *three out of five neighbors*: Three final judgment scores of the five most similar database trials (compared to the evaluation trial in terms of movement kinematics) were averaged to give the score for the evaluation trial. Thereby, the highest and lowest final judgment scores were omitted from this calculation. 4) Model variant #4 *recurrent neural network*: The information on the specific pattern of the neural network concerning the kinematics data was used as input.

Model Prediction

For the training and evaluation trials, $n = 20$ trials for each skill were randomly selected, to get an amount of approximately two-thirds for the model's database and an amount of approximately one-third for the training database (Table 9.1). We needed to ensure that the database trials and evaluation trials did not differ in average scoring, and were thus suitable for prediction modeling. Therefore, the mean scores of the database trials were compared to the mean scores of the evaluation trials by calculating the Wilcoxon

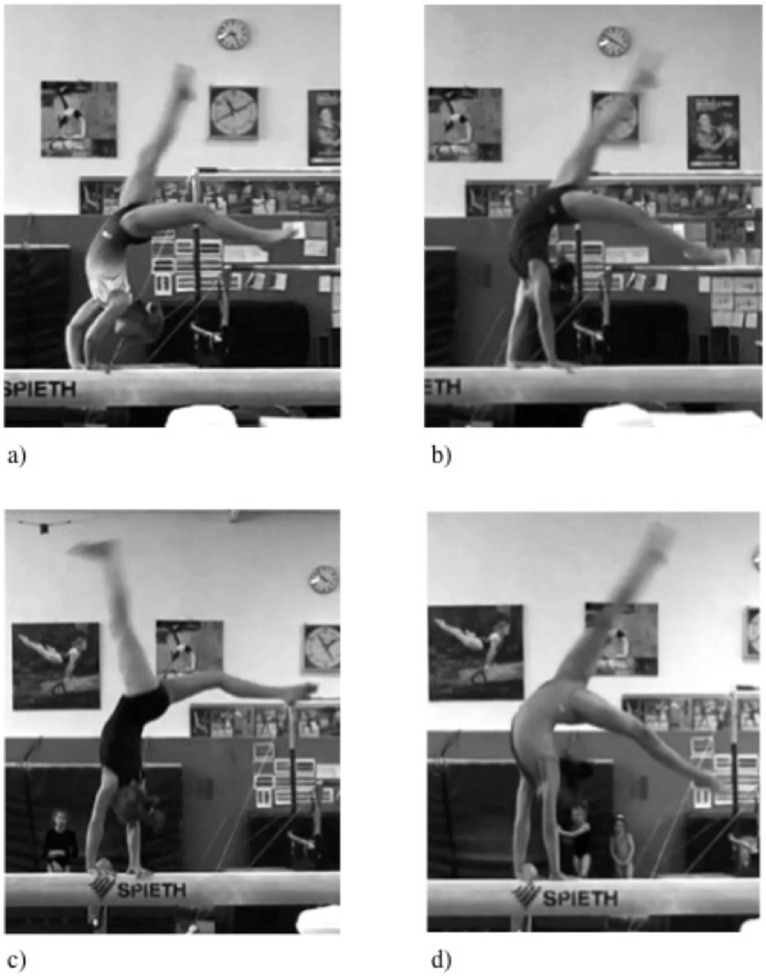

FIGURE 9.1: The picture sequence displays structural similarity. Body positions a) and b) have a high structural similarity, whereas the other body positions c) and d) have a lower structural similarity.

rank-sum test, because of their abnormal distribution. For the model variant #4 *recurrent neural network*, the training was as follows: The weights of the networks were initialized randomly, and thus, the output of the network was useless in the beginning. The weights were adjusted to give better outputs. The adjustment was performed using an algorithm called the backpropagation algorithm. When using backpropagation, an error was calculated at the output layer, whereby that error was the difference between the values the network created and the values the network should have created. Using that error and the targets, the change in the weights was calculated so that the output approached the desired values. To estimate the prediction performance of each model variant, we calculated Spearman's rank correlation coefficients between the predicted and original scores of the evaluation trials for all model variants and all apparatuses (Glöckner, Heinen, Johnson, & Raab, 2012).

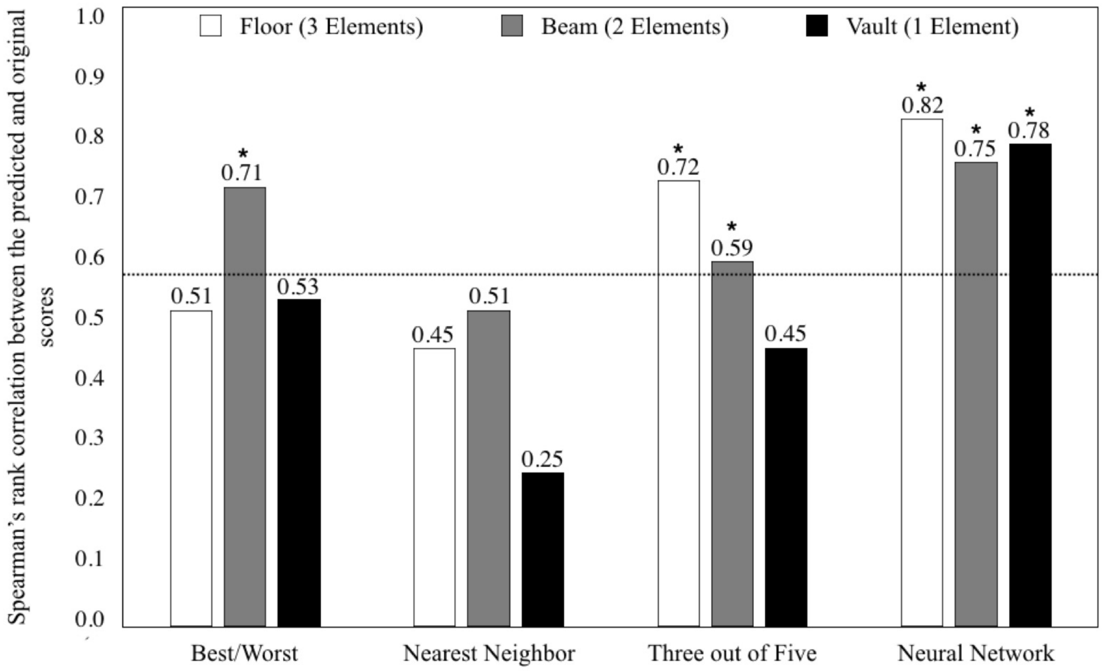

FIGURE 9.2: Spearman's rank correlation between the predicted and original scores for the five different model variants and the three different apparatuses. *Note*: The dotted line indicated the critical r-value (textitp = .05)

Results

The Shapiro-Wilk test of normality was used to investigate whether the judgment scores (predicted and original scores) were approximately normally distributed. Most of the scores were abnormally distributed. Therefore, the Wilcoxon rank-sum test was used, to ensure that the database trials and evaluation trials did not differ in average scoring and revealed no significant differences between scores for any of the three apparatuses. Table 9.1 presents the mean original scores of the database trials and evaluation trials for the three apparatuses as well as for the results of the Wilcoxon rank-sum test.

The prediction performance of each model variant was estimated by calculating the Spearman's rank correlation coefficients between the predicted scores for the model and the original scores by the judges for each of the three gymnastics skills. The significance level was defined a priori as 5 %. A one-tailed bivariate correlation with $N = 20$ cases, a type-I error probability of $p = .05$ and a Bonferroni correction could be calculated with a value of $r_{crit} = .57$. Thus, for a correlation coefficient to become statistically significant, its value had to be larger than r_{crit}.(Figure 9.2). In addition, the predicted mean scores of all evaluations were calculated and compared to the original mean scores of all evaluation trials using the Wilcoxon rank-sum test; this was done for all model variants and all apparatuses (Table 9.1). Figure 9.2 presents the calculated Spearman's

Appa-ratus	Model Variant Best/ Worst	Z	p	Nearest Neighbor	Z	p	Three out of Five	Z	p	Recurrent Neural Network	Z	p
Floor	3.73 ± 1.22	1.47	.140	3.46 ± 1.75	1.85	.063	3.84 ± 1.69	0.34	.735	3.70 ± 0.79	2.18	.029*
Beam	3.87 ± 2.38	0.85	.394	3.65 ± 2.13	0.31	.756	3.91 ± 2.11	0.89	.372	4.30 ± 1.00	0.86	.390
Vault	3.30 ± 2.10	0.06	.947	3.34 ± 1.89	0.07	.946	3.53 ± 1.75	0.39	.695	3.10 ± 1.11	1.88	.060

Table 9.1: The number of trials used to generate the model database, the number of trials used to evaluate the model, inter-rater reliability for the judges' scores (ICC), mean scores of the database trials (standard deviation) and evaluation trials, as well as the results of the Wilcoxon rank-sum test. *Note*: The inter-rater reliability was calculated for all three judges. Scores were assigned between one and six points, according to the judging guidelines of the German Gymnastics Federation for talented young gymnasts (FIG, 2017).

rank-sum correlation coefficients per model variant and per apparatus. Table 9.1 highlights the mean original scores of the judges for the evaluation trials as well as the mean scores of the evaluation trials which were predicted by the different model variants.

First, six out of twelve correlation coefficients reached statistical significance with regard to a critical value of $r = .57$ (with a Bonferroni corrected $p < .004$), thereby indicating a significant relationship between the predicted scores and original scores of the evaluation trials (Figure 9.2). Only the model variant #4 *recurrent neural network* exhibited a significant relationship between the predicted and original scores for all apparatuses: floor, $r_s(56) = 0.82$, $p < .001$; beam, $r_s(49) = 0.75$, $p < .001$; and vault, $r_s(56) =$, 0.78, $p < .001$. The model variant #2 *nearest neighbor* did not exhibit a significant relationship between the predicted and original scores for any of the apparatuses: floor, $r_s(20) = 0.45$, p .< .044; beam, $r_s(20) = 0.51$, $p < .023$; and vault, $r_s(20) = 0.25$, p < .292. The model variant #1 *best/worst* only showed a significant relationship for the beam, $r_s(20) = 0.71$, $p < .001$, whereas the results were not significant for the floor, $r_s(20) = 0.51$, $p = .021$, and the vault, $r_s(20) = 0.53$, $p = .016$. For the model variant #3 *three out of five neighbors*, only the floor, $r_s(20) = 0.72$, $p < .001$, and beam, $r_s(20)$ $= 0.59$, $p = .003$, showed a significant correlation, whereas the correlation for the vault, $r_s(20) = 0.45$, $p = .48$, was not significant.

Discussion

The aim of this study was to develop and evaluate a model-based approach for judging gymnastics performance, based on quantitative kinematic data of the performed skills. Four different model variants were compiled, in order to predict the judgment scores on the basis of the kinematic information. The results showed a significant relationship

Apparatus	Database Trials	Evaluation Trials	ICC_{all}	Mean Score of Database Trials	Mean Score of Evaluation Trials	*Z*	*p*
Floor	38	20	0.75	3.94 ± 1.90	3.99 ± 1.99	0.08	.818
Beam	37	20	0.85	3.94 ± 2.23	3.52 ± 2.31	1.24	.216
Vault	37	20	0.83	3.35 ± 1.90	3.42 ± 1.81	0.23	.933

Table 9.2: Comparison of the judges' original scores for the evaluation trials and the predicted scores from the different model variants (mean $\neq$ standard deviations), as well as the results of the Wilcoxon rank-sum test. *Note*: $*$ denotes a statistical difference between the original and the predicted scores.

between the predicted scores and the original scores for six of the twelve gymnastic skill-model combinations. The different model variants yielded different prediction performances in general, over all skills, and also in terms of the different skills. It was assumed that gymnastics skill performances with similar kinematic patterns reveal similar evaluation scores. Similarity is thereby defined in terms of the information about the time courses of the main body angles. These are important characteristics, because they describe gymnastics skills in a holistic way, and other kinematic characteristics can easily be computed from these values (Enoka, 2002). Our approach is similar to the approach of Díaz-Pereira et al. (2014) who assessed velocity covariance trajectories rather than the angle values. The authors showed that the covariant velocity trajectories contain information about the spatio-temporal aspects of a particular motion, for extraction of the quality differences between movements. However, this algorithm may work for simple movements, but it is not suitable for complex movements. By taking time-courses of the body landmarks as well as the resulting angle courses, the approach describes another tool that is capable of recognizing and scoring the movement quality of complex gymnastic skills.

It is assumed that the kinematic pattern of a gymnastics skill contains the relevant information for an accurate judgment of performance (Blake & Shiffrar, 2007). Several researchers have illustrated that judges make their evaluation not only on the basis of kinematic similarity but also by taking some more subtle features into account, that are not directly captured via movement kinematics (Bertenthal, Proffitt, Spetner, & Thomas, 1985; Cutting et al., 1978). Those subtle features may lead to bias in the judgment process, because their perception and encoding may be influenced by social cognition, information processing or memory processes (Plessner & Haar, 2006). These assumptions lead to the idea of making the judgment in artistic gymnastics more objective, by using technical methodologies that predict judgment scores on the basis of kinematic variables. The study is the first approach in this direction, and offers interesting results.

There are several aspects of the results that should be discussed in terms of this assumption. Prediction models that take into account only one trial model variant #2 *nearest neighbor*, two trial model variants #1 *best/worst* or five trial model variants #3 *three out of five neighbors* were less accurate in their score prediction than the neural network approach. This was particularly the case for the gymnastics movement, handspring on vault, which has a rather short duration. This result may be explained by two points. First, because of the short duration of the skill, the time available to receive the relevant kinematic movement information is limited, whereas the observer has more time available for a routine during the floor exercise, which incorporates three gymnastics skills in a row. Secondly, because the vault always comprises one element, the complexity and velocity of the movement are higher than for the other elements, which also could lead to less accurate results. Because of the demands of the vault, in general, this apparatus should be first to benefit from a kinematic assessment, compared to other apparatuses. It could be interesting to compare the kinematic computer-based judgment results with the judgment results of gymnastics experts acquired from slow-motion videos.

Generally, the prediction performance was the best for the neural network model. The trials in which the network gave a different score than the majority of the judges, was interesting at this point. An indicator that a neural network has started learning and has understood the given data is that it does not just give the same score for every test input, but also gives high scores for the same scores where the judges gave high scores, low scores for other trials accordingly. This could mean that the network is being more objective and thus comes to a different conclusion, compared to that of the judges. Assuming that the scores from the network are accurate, an interesting use case comes to mind. Since the network can basically create scores instantly, it could be embedded in a program that films gymnasts and then immediately gives feedback after the scores have been assigned, indicating where the kinematics variables should be changed. Thus, the approach could be applied in competitions, perhaps not as an alternative approach to a human evaluation but for complementing the actual scoring procedure. The recording and digitizing of the performance were performed using a semiautomatic system. There are hardware options available that are more precise in their tracking of the relevant kinematic features, which release a higher amount of kinematics variables and are much faster in their digitizing process. For instance, there is direct three-dimensional silhouette tracking software available that extracts movement kinematics of the human body by means of high-performance silhouette tracking algorithms (Colyer, Evans, Cosker, & Salo, 2018).

Another advantage is that the approach can be easily applied to data from different motion capture systems because the algorithm can be adapted to various kinematic variables. By having a larger dataset and different kinematic variables, one could compare a larger number of different models and thereby find the most precise one. This leads us to the limitation of the study where one specific aspect should be highlighted. First, two of the models were a combined reference-based and nearest-neighbor averaging approach or a purely reference-based approach. By taking into account not only the kinematic information, but also the scoring information for the model approaches, one uses the scoring information to predict the score. By labeling the kinematic information as a stable factor and the scoring information as a variable (changeable through different raters) factor that contains human bias, the model approaches differ in their independence from the human evaluation bias. The variability of the variable factor changes, and it should increase by a larger number of trials and scores per trial. The better prediction of the prediction model neural network could thereby arise from the large number of trials in the model but also from the lower bias originating from the variable factor.

Furthermore, it would be interesting to investigate not only whether models containing different variables and a different number of variables tracked from different motion capture systems leads to a different score prediction, but also which variables are most precise in terms of predicting the score. For example, if a model that only takes the hip angle into account leads to the same results as a model that takes four or five time courses of body angles into account, but this is not the case for a model on the basis of the knee angle, then it might be assumed that the hip angle is more relevant for evaluating gymnastics skills than the knee angle. Furthermore, the approach should be tested by taking three-dimensional instead of two-dimensional kinematics data, and by testing the models with skill realization data at a broader base of expertise level and skills that improve the procedure of achieving the judgment scores.

One factor that should always be kept in mind when dealing with computer-based methodologies is the psychological aspects of judging human behavior. It can be assumed that being judged by a computer or a human being leads to behavioral differences in skill execution. It may be assumed that each performance has a certain emotional expression that could hardly be captured by computer-based technologies. On the other hand, there is the question of trust in computer-based technologies and their error rate. They are more objective than human beings, but there are many aspects of motion capturing or of the algorithm that could lead to errors.

Conclusion

Overall, the approach utilized in this study to predict the evaluation scores of different gymnastics skills, using a combined reference-based and nearest-neighbor averaging approach, is a novel and important topic, as the FIG is attempting to implement a judging support system based on movement data. The study revealed the first interesting results that offer both practical applications and further research questions, to complement the judging procedure in gymnastics competition, or similar areas of sport, with technical methodologies.

Chapter 10

Discussion

This thesis aimed, in a first step, to kinematically analyze selected prototypical gymnastics skills (Mack et al., 2018; Mack et al., in press), in order to investigate, in a second step, the underlying perceptual-cognitive mechanism when perceiving and evaluating those skills (Mack, 2019, Mack et al., 2019, Mack et al., submitted. In the analysis of kinematics, it was focused on the classification and structuring of temporal-spatial continuous kinematic movement patterns (Mack et al., 2018; Mack et al., in press). In the investigation of the perceptual-cognitive mechanism, it was focused on the relationship between skill kinematics and mental representation (Mack et al., submitted), gaze pattern (Mack, 2019) and the estimated performance quality of the evaluators (Mack et al., 2019). Five consecutive studies were conducted, which will be discussed concerning their key findings, their theoretical background, their method, and their practical implications.

10.1 Key Findings

Motor Behavior

The analysis of the kinematics was exemplarily realized with the gymnastics skill handspring on vault (Mack et al., 2018), as well as a floor routine comprising the three elements, round off, back handspring and back layout somersault, in sequence (Mack et al., in press). The results revealed that for near experts, three to four prototypical movement patterns could be identified; for all, the single element on vault, the entire floor routine, and also the three particular skills of the floor routine separately. The different prototypes could be differentiated by certain variant and invariant characteristics, such as the time course of the different body angles and their coefficient of variation.

With regard to the distribution of the trials of one person to the prototypes, not all trials from one gymnast were associated with one particular prototype. The interactions of the complex system could explain this variability. The interactions might change, depending on the constraints provided by the gymnast, the environment, or the task (Bernstein, 1967; Davids et al., 2014; Higgins, 1977). But a general tendency was discernible, which indicates that the movement pattern of a gymnast remains similar over a majority of the trials (Troje, 2002).

When comparing the three to four prototypical movement patterns using the judgment criteria defined in the Code of Points (FIG, 2017), the prototypes could be distinguished regarding the way they match those criteria. For instance, concerning the skill handspring on vault, guidelines are defined in the Code of Points, regarding the movement pattern of the hip, the shoulder, and the knees. The trials, which were included in two of the four prototypes, might thereby be scored higher than the trials of the other two prototypes. For the floor routine, where three gymnastics skills were done in a row, a statistically significant relationship, in terms of the distribution to the prototypes, could, to some extent, be found . This means, for example, that gymnasts who execute the round off in a similar way, perform the back handspring similarly too. Therefore, it might be assumed that certain manifestations of movement patterns (for example, the time-space pattern of the knee, shoulder, or hip joint) can also be detected in the subsequent skills.

The characteristics of those two exemplarily investigated events (vault and floor) differ starkly. Firstly, with regard to the number of elements and thereby also in terms of the temporal duration of the skills (one on vault and three on the floor), and secondly, with regard to the local variability of the skills. On vault, the place where the skills can be performed is defined very precisely through the construction circumstances of the event. The hands have to be in a particular position on the vault. On the floor, the place where the skill should be performed is also defined and limited, but it is up to the gymnast to decide how to spread the three skills and the approach over the available area. Thirdly, the characteristics differ in terms of the velocity and the complexity. Because the vault always comprises one element, the velocity (and often also the complexity) of the movements are higher than for the other events. Nevertheless, similar results could be found concerning the number of prototypes, which speaks in favor of this approach.

The results of the model-based approach to judging gymnastics performance on the basis of quantitative kinematic data showed a significant relationship between the predicted scores and the original scores for six of the twelve gymnastics skill-model combinations (Mack et al., 2019). The different model variants yielded different prediction performances in general, overall skills, and also in terms of the different skills. Prediction models that only take one trial into account (nearest neighbor), two trials (best/worst), or five trials (three of five) were less accurate in their score prediction than the neural network approach, which takes all trials into account. This was especially the case for the shortest movement, the handspring on vault, which could be explained by two points: First, the limited time available for receiving the relevant information and second, the complexity and velocity of the vault elements. Concerning the argument of the limited time available, one can easily refer to the results of the study that analyze the floor routine kinematically (Mack et al., in press). It could be assumed that specific manifestations of movement patterns (for example, the time-space pattern of the knee, shoulder, or hip joint) can also be detected in the subsequent skills. Therefore, because it is probable that the consecutive skills were similar and also have a similar performance quality, there is more time to perceive the relevant movement information. Overall, because the FIG is attempting to implement a judging support system based on movement data, the study revealed interesting results regarding objective computer-based approaches for evaluating gymnastics movement, based on kinematics data.

Perception and Cognition

Mental Representation of Movement Quality

To answer the question that emerges, of whether skills with a similar kinematics pattern (associated with one particular prototype) were seen as more similar than skills with a less similar kinematic pattern (associated with different prototypes), both the relationship between the perceived movement quality of a gymnastics skill and its kinematic pattern and the influence of expertise in this relationship were investigated (Mack et al., submitted). The comparison of the resulting cluster solution of the perceived movement quality of the three experimental groups (motor experts, visual experts, and novices) with the kinematics pattern revealed no significant correlation. It is possible that either the perceived movement quality was not related to the holistic kinematics pattern or that the evaluation of movement quality was not only based on a consistent influence of the kinematics features. The actual research results speak in favor of the second explanation, indicating that judges not only base their rating on the kinematics features but also on other subtle features, such as dress, facial expression, mental capacities, personal experiences and expertise, as well as the point of view and the order of presentation of the gymnasts, all of which play a role in the evaluation process (Ansorge et al., 1978; de Bruine, 2006; Heinen et al., 2013; Pizzera, 2012; Pizzera & Raab, 2012; Plessner & Haar, 2006; Plessner & Schallis, 2005; Ste-Marie, 1999). It could also be assumed that particular kinematics features have different weighting in the evaluation process. For example, the landing position might influence the judgment more than the starting position. To gain insight into this hypothesis, methodologies such as neural networks could be taken into account.

A comparison of the cluster solution of the three experimental groups revealed a significant correlation between the dendrograms of the visual experts and the motor experts, as well as between the dendrograms of the visual experts and the novices, but no significant correlation between the dendrograms of the motor experts and novices. Similarly to the above mentioned information, this might be explained by a different weighting of different kinematics aspects, dependent on the level of expertise. These findings are in line with research results concerning the visual search strategy when evaluating gymnastics movements. Here, it could be shown that a different level of gymnastics expertise led to focusing on different information sources on the visual display (Bard et al., 1980; Del Campo & Gracia, 2018; Pizzera et al., 2018). The differences became apparent in the number of clusters and in the closeness of the different trials and clusters. Two clusters could be identified for both the motor experts and the visual experts, and three clusters could be determined for the novices. Thereby, the two distinct clusters of the dendrograms of the visual and motor experts were less similar than the three different clusters of the dendrograms of the novices. This different variance in the evaluation of movement quality might have occurred because of a differently detailed mental representation of the kinematic pattern or because of a different concept regarding the qualitative evaluation regularities of movement patterns and particular movement characteristics. It could also be demonstrated that overall, visual experts rated different gymnastics floor performances worse and less accurately than motor experts and novices (Mack, 2019). Those different results reflect the actual state of research, where a variant picture concerning the evaluation results could be found(Čuk, 2015; Ste-Marie, 2003). The worse rating by visual experts could be explained by their higher level of visual experience, leading to a recognition of more and smaller mistakes. This is in line with the main effect for format, with regard to the evaluation, in which videos in the stick-figure format were rated higher than videos in the original format. The abundance of minor movement errors could account for this effect. The results for judgment accuracy were more counterintuitive. The evaluation task used here, which was not similar to the evaluation procedure in a competition, might have led to bad habituation of the visual experts to the evaluation instructions given.

Gaze Pattern

Interestingly, against the assumptions that expertise leads to a different gaze pattern (Bard et al., 1980; Del Campo & Gracia, 2018; Pizzera et al., 2018), no main effect could be found for expertise, with respect to the entire floor-routine, but also to the different skills (Mack, 2019). Visual experts, motor experts, and novices showed a similar gaze pattern. Regarding the gaze pattern of the entire floor routine, a main effect for format could be found for the variable average fixation duration and the variable number of fixations. Performances that were presented in the stick-figure format led to a higher average fixation duration and a lower number of fixations then performances that were presented in the original format. Because it is supposed that a longer fixation duration leads to more detailed information processing, by allowing the extraction of a higher volume of information (Mann et al., 2007; Williams et al., 1999), the longer average fixation duration for the stick-figure format could be explained by the lower volume of information presented.

Regarding the gaze pattern of the different temporal-spatial windows, there was a change in the direction of the significant main effect for the variable number of fixations. For the original format, the results changed from a higher number of fixations regarding the analysis of the entire floor routine to a lower number, regarding the analysis of the temporal-spatial windows. This change might lead to the assumption that in the stick-figure format, the participants were not occupied with extracting and processing the information, which was shown in another temporal-spatial window, and thereby might have focused more on the temporal-spatial window of the shown skill. Significant interaction effects could be found for the variable summarized fixation duration and number of fixations. Thereby, the stick-figure condition led to a higher number of fixations for the back handspring and the back layout somersault and a longer summarized fixation duration for the back layout somersault. This interaction might lead to the assumption that the last skill plays a special role in the judgment of a floor routine.

10.2 Theoretical Background

The perceived performance quality plays an integral role in gymnastics for judges, coaches, athletes, and spectators. While judges (and to a certain degree also gymnasts and coaches) observe movements regarding criteria defined in the code of points, and base their judgment on those regularizes, even spectators without knowledge of those guidelines came to a similar judgment of presented skills (Čuk, 2015; Ste-Marie, 2003). The complex process of observation and evaluation of gymnastics skills can be located within the interplay between action and perception and the theoretical framework that perception and action are represented and saved in the same area of the brain. (Dosseville et al., 2014; Prinz, 1997). When aiming to analyze action, one thereby should not leave perception out of the considerations, and the other way round, when seeking to analyze perception, one should not ignore the action. This was the theoretical guiding principle of this thesis. Because the evaluation guidelines describe kinematic characteristics, the action was thereby taken as the quantitative and qualitative analysis of biomechanical variables of the motion. The perception was thereby the analysis of the gaze pattern when observing movement, the cognitive representation and information processing of the motion, and the resulting evaluation score.

Motor Behavior

Concerning action, in most cases, both literature and the practice dealing with gymnastics movements, address the movement pattern ("technique") of gymnastics skills and the realization thereof, by describing the relative position and orientation of body segments as well as their coordination during the performance of a skill (Lees, 2002). This is strongly linked to the theoretical model of complex systems, which is the reason why this model was taken as a theoretical background for investigating action. How this theoretical model is associated with the investigation of perception will be discussed later on. The theory describes every movement as a complex system consisting of many components that interact among themselves, and with the environment as a whole (Davids et al., 2014). The high number of components leads to a high range of degrees of freedom in the motor system. As a result, motion in general (and therefore, also gymnastics skills) can be performed in many different ways (Bernstein, 1967; Latash et al., 2002). This interaction and, therefore also the variation, is influenced by constraints (provided by the task, the environment, or psychological factors). Literature suggests that within this space of possible movement patterns, there are detectable and characterizable prototypical movement patterns that are similar in terms of their coordination dynamics (Jaitner et al., 2001; Schöllhorn et al., 2009) and that those prototypes can be structured, segmented and classified (Endres et al., 2012; Ren et al., 2005; Salamah et al., 2015). This theoretical assumption is strengthened by the results of the studies presented here, where for different skills, a particular number of prototypes could be detected (Mack et al., 2018; Mack et al., submitted). Literature further suggests that those prototypical movement patterns can be changed by learning and training(Davids et al., 2008). Through a similar level of gymnasts' skills and the lack of a longitudinal investigation, the studies presented here could not prove this theoretical assumption.

Perception and Cognition

Mental Representation of Movement Quality

When now switching to the perception, in gymnastics, the ability to identify relevant information in the movement and link them to prior knowledge, in particular, as well as the expertise illustrated thereby, seems to be pertinent (Gegenfurtner et al., 2011; Mann et al., 2007; Williams & Ericcson, 2005). It is thereby suggested that those mental structures are linked with the structure of the movement kinematics, and that the mental representation resembles the action in terms of the structural similarity (Schack's architecture model: Schack, 2004; Schack & Ritter, 2009; Schack, 2010). In sports such as artistic gymnastics or dancing, not only the anticipated effects of a movement are of importance. It is also essential to evaluate the movement quality of skills as a whole. However, it could be shown that the cluster solution of the perceived movement quality (motor experts, visual experts, and novices) did not correlate significantly with the kinematics pattern (Mack et al., submitted). The reason for this might be that particular kinematics features were taken into account with a different weighting when evaluating the quality of a particular movement.

Motion Perception

On the level of how motion itself is processed, it is suggested, that perception is a holistic process in which relative motion, based on kinematics cues over time and space, were used (Cutting et al., 1978; Runeson & Frykholm, 1983; Troje, Westhoff & Lavrov, 2005). It is further suggested that several indices of gaze (the location, the number, and the duration of fixations) influence the process of information pick-up and processing and are, on the other hand, influenced through the prior knowledge and the mental representation of the movement. The other way round, the presented information on the visual display should also influence those gaze indices and thereby the processing. It might be assumed, that also for the perception and evaluation of gymnastics skills, the kinematics are of mere importance, because by using deterministic models, it could be shown that the kinematics predictors affecting the final score out of a large set of potential predictors (for example: Farana & Vaverka, 2012; Farana, Zahradník, Uchytil, & Jandaka, 2015; Takei, 1990; 1992, 1998, 2007; Takei et al., 2000).

One method of investigating the influence of kinematics in the evaluation process is the use of point-light or stick figure displays (Barclay et al., 1978; Cutting & Kozlowski, 1977; Cutting et al., 1978; Dittrich, 1993; Stevenage et al., 1999; Troje, 2002; Troje et al., 2005). By merely displaying kinematics information, this might influence the information processing workload. The results presented here confirm this assumption by showing a higher average fixation duration and a lower number of fixations for performances presented in the stick figure format than for performances presented in original format (Mack, 2019). This might be because a lower processing load is required to process the stick-figure videos. The format did not influence the judgment performance.

Concerning the particular skills, different gaze patterns between videos of the two formats could be found for the last skill. This leads to the assumption that the last skill plays a particular role in the judgment. This is in line with a psychological effect called the serial-position effect, which is the tendency to recall the first and the last items of a series best and the items between worst. Such serial position effects have already been identified for judgment in sporting events, where it was found that a benefit existed when performing later in the sequence (de Bruine, 2006)

10.3 Method

Motor Behavior

Most of the research concerning the kinematic analysis of gymnastics skills deals with selected variables. Because humans perceive movement holistically , the goal was to develop a method to analyze a front handspring on vault as well as a gymnastics skill routine (round off, back handspring, backward layout somersault) in a holistic fashion. Therefore the digitized video sequences of 60 handspring trials and 58 floor routine trials (round off, back handspring, backward layout somersault) from ten near-expert female gymnasts were used for kinematic analysis. Time courses of six joints were analyzed and structured in prototypes, using hierarchical cluster analysis. Besides this, the coefficients of variation were calculated for each prototypical movement pattern (Mack et al., 2018; Mack et al., in press). To investigate the relationship between the three gymnastics skills of the floor routine, Cramer's V correlations between the different cluster were conducted (Mack et al., in press). Overall, the approach utilized in those studies allowed to identify structural characteristics of movement patterns of complex gymnastics skills. This method seems not only appropriate for one particular skill but also a wide range of gymnastics skills. In both studies, the skill executions used were conducted by near-expert female gymnasts. Thus, on the one side, the results of the different skills could be compared, and the conclusion of the possibility to transmit the approach on a wide range of gymnastics skills could be drawn. On the other hand, it is unclear whether those results could be transferred to gymnasts with a different level of expertise and also whether the distribution of the trials to the different prototypes changes with training. One might assume that the movements of experts are more similar than for novices and, therefore, that the distribution to the different clusters would be less widespread than for novices. It is unclear whether expertise leads to a different amount of prototypes compared to novices. On the one side, there is an optimal kinematic pattern prescribed in the code of points, and the differences between the performances are often small (GymnasticsResultsCom, 2012), but on the other side, the research could show that even the movement characteristics of experts differ strongly.

Mental Representation of Movement

To answer the question, whether skills with a similar kinematics pattern (associated with one particular prototype) were seen as more similar than skills with a less similar kinematic pattern (related to different prototypes), the relationship between the perceived movement quality of a gymnastics skill and its kinematic pattern was investigated, as was the influence of expertise in this relationship (Mack et al., 2019). By means of a paired comparison test and an ongoing cluster analysis, the pattern of the perceived movement quality of the gymnastics skills (handstand, back handspring on the balance beam) was determined for each experimental group (n = 10 visual experts, n = 10 motor experts and n = 10 novices), analyzed and compared with the pattern of the perceived movement quality pattern of the other groups, as well as with the kinematic pattern of the investigated skills. The study builds, conceptually and methodologically, on previous research that has identified links between the mental representation of a movement and the movement kinematics (Schack, 2003). The methodology used here was an initial approach to comparing the temporal-spatial continuous structure of the kinematic pattern of human movement with the mental representation of its qualitative evaluation. Through the hierarchical cluster analysis, the resulting structure of the perceived movement quality of each experimental group, as well as the kinematics pattern involved, was displayed and visualized, inspired by approaches such as the SDA method (Lander & Lange, 1996), the SDA-M method (Schack, 2004) or the Repertory Grid Technique (Fransella, Bell, & Bannister, 2004). Nevertheless, it could not be determined how the relationship is influenced by a different weighting of different kinematics aspects.

Motion Perception

To gain insight into the perceptual-cognitive processes when evaluating gymnastics movement and more specifically the role of kinematics, a combination of methods that include both the performer's gaze pattern (via eye-tracking) and their performance judgment (via ratings on a six-point Likert scale) was employed for gymnastics performances which were shown as videos in original or stick-figure format (Mack, 2019). The gaze pattern was thereby analyzed for the whole floor routine as well as for the particular spatial-temporal windows of the three motor skills. To the best of my knowledge, the method used herein is the first that investigated judgment and gaze patterns concerning the influence of essential kinematics information and the varying importance of the particular skills in a sequence. To investigate the gaze pattern concerning the different skills, it was focused separately at the gaze patterns, which were shown at the specific time-space window of the three different skills. For instance, for the round off skill, only the gaze pattern, which was shown at the spatial window of the screen, and the particular time-window, where the round off was performed, was extracted. This method does not consider whether, in those windows, the gymnast or the environment was focused upon. Because the size of the visual field which was occupied by the gymnast was not very large, and the velocity of the motion was very fast, the probability of losing a lot of information in terms of the gaze pattern, would have been very high if analyzing only the gaze-data of the fixations on the gymnasts.

To evaluate the relationship between the movement kinematics and the judges' scores, a model-based approach to judging gymnastics performance, based on quantitative kinematic data, was developed (Mack et al., 2019). Four different model variants, calculated by a multivariate exploratory approach and the Recurrent Neural Network method, were used. The complete dataset consisted of movement kinematic data and judgment scores of a total of N = 173 trials of three different skills and routines from women's artistic gymnastics, on a near-expert level. With the five models, the bandwidth of possibilities of judging was covered, from taking the order of the judgments to taking five judgments and removing the best and the worst, to taking only the kinematic pattern. This combined reference-based and nearest-neighbor averaging approach was thereby a novel and important method, as the FIG is attempting to implement a judging support system, based on movement data. On the other side, there are limitations to the approach, which should be mentioned. Two of the models were a combined reference-based and nearest-neighbor averaging approach or a purely reference-based approach that takes not only the kinematic information, but also the scoring information, into account, which means that one used the scoring information to predict the score. Contrary to the kinematic information, which could be labeled as a stable factor, the scoring information is variable. The better prediction of the model neural network could thereby arise from a lower bias originating from the variable factor, but also from the large number of trials in the model.

The approach is similar to the approach of Díaz-Pereira, Gomez-Conde, Escalona, & Olivieri (2014), who assessed velocity covariance trajectories rather than angle values. Thereby it could be shown that the covariant velocity trajectories contain information about the spatio-temporal aspects of a particular motion, which can be used to extract the quality differences between movements. By taking time-courses of the body landmarks as well as the resulting angle courses, the approach describes another tool that is capable of recognizing and scoring the movement quality of complex gymnastic skills. Another advantage is that the approach can be easily applied to data from different motion capture systems, because the algorithm can be adapted to various kinematic variables. By having a larger dataset and different kinematic variables, one could compare a larger number of different models, and thereby find the most precise one. Concerning the neural network model, the trials in which the network resulted in a different score than the majority of the judges, could be interesting for further investigation. A study could be conducted, to determine which variables lead to those deviations. Assuming that the scores from the network are accurate, and since the network can create scores instantly, it could be embedded into a program that films gymnasts and then immediately gives feedback after the scores have been assigned, to indicate where the kinematics variables should be changed.

Skill Dataset and Movement Analysis

The movement data set used consists of a total of N = 173 videos of three different skills from women's artistic gymnastics (round off - back handspring - back layout somersault on floor, n1 = 58; back handspring from handstand position on balance beam, n2 = 57; handspring on vault, n3 = 57). All skills were performed by the same ten high-level, female gymnasts, whereby every athlete performed six to seven executions of each skill. The fact that the skills of the different events were performed by the same athletes leads to a good comparability of the results of the different events (Mack et al., 2018; Mack et al., 2019; Mack et al., in press). On the other hand, the similar level of the gymnasts allowed no transfer of the results to gymnasts with a higher or lower expertise level. In further studies, the influence of expertise of the athletes could be investigated. Because the studies presented here were all cross-sectional studies, the influence of training through longitudinal studies would be of further interest.

Furthermore, it is worth mentioning that the videos were cut at the beginning and at the end of the skill, so that only the movement itself was shown, and not the starting and landing positions. This leads to a reduction of bias on the object of interest, namely the movement, as a result of the starting or landing, but on the other hand, also to less transfer of the results to the practice. The kinematics data used was confined to a two-dimensional time series of the calculated joint angles of the knee, hip, and shoulder (in each case, for the left and right body side). First, the two-dimensionality and second, disregarding several body angles (for example elbow and ankle leads to an incomplete picture of the kinematic pattern). Anyway, it should be stated that the performances were videotaped orthogonally to the direction of movement, simulating the judges perspective, and that elbow and ankle joints were excluded from the analysis, because they were biased through the two-dimensional measurements. In further studies, not only the use of three-dimensional data but also time warping approaches could be taken into consideration, to account for the different length of movements.

Visual-motor Expertise

Experimental groups with different levels of gymnastics expertise (visual experts, motor experts, and novices) were used to measure the influence of expertise on cognition and perception (Mack, 2019; Mack et al., submitted). The motor experts group consisted of persons who had been active in gymnastics as athletes for a minimum of three years and were still active at the time of the investigation. The visual experts group included persons who had been active in gymnastics as judges or coaches for a minimum of three years and were still active at the time of the investigation. The novices group included persons who were neither active, nor had they ever been active in gymnastics, either as athletes, judges, or coaches. Here, it is worth mentioning that most of the participants had been active longer than presumed, but that neither the middle expertise level of the motor experts nor the expertise level of the visual experts amounted to high-level expertise. One must further note that, in gymnastics, visual and motor expertise is mostly confounded by persons being athlete and judge and/or coach at the same time. We attempted to consider this aspect by excluding the participants with a similar level of visual and motor expertise and primarily taking participants with a focus on either judging and coaching or those who were active as an athlete in training and competition.

Judgment

To operationalize performance quality, a six-point Likert scale, according to the judgment guidelines of the German Gymnastics Federation for young gymnastics talents (DTB, 2001), was used (Mack et al., 2018; Mack et al., in press, Mack, 2019; Mack et al., 2019). It was chosen to do so, in order to use a tool that is on the one side already validated in practice, and on the other side, to provide a procedure with which novices are also familiar (because of the similarity to the grading system in the school context). But one should further consider that this forced judges to evaluate performances differently than they are used to doing in competitions, which could have an influence on the results of both the eye-tracking experiment (Mack, 2019) and the model approach (Mack et al., 2019). It would be interesting to replicate the studies with performance scores resulting from a judging procedure used in competitions.

10.4 Practical Application

Training Process

From the perspective of the dynamic system theory, the emergence of different kinematics prototypes (Mack et al., 2018; Mack et al., in press) leads to some reflections concerning their potential role in terms of motor learning. Because it is assumed that through learning the entire layout of the coordination dynamics changes (Bernstein, 1967), the contribution of skill performances to prototypes might reveal the point at which motor behavior starts changing (maybe as an indicator of critical learning phases). It might be stated, that individuality in training in terms of an optimal organization of the complex functional movement system, is an essential factor in learning and improving gymnastics skills. Therefore, the approach might be used as diagnostic tool. The skill execution of one athlete could be categorized by a comparison of the performance with the different prototypes, and stations in the learning process could thereby be detected, and helpful instructions for the training process could be selected.

For instance, if the contribution of the performances of a certain skill of a certain athlete is stable in terms of one particular prototype, the movement pattern is also stable. If the contribution is not stable, not even the movement pattern is stable. If, over time, the contribution pattern changes from stable to unstable, or from unstable to stable, motor behavior starts changing or starts becoming stable. In the case of switching from stable to unstable, it might be essential to focus the training on this particular skill, in order to automate a specific movement behavior. In training, the question of finding the optimal point of time to proceed to learn more complex versions of a particular skill, for example, a double somersault instead of a single somersault, is also appreciable. The perfect point in time might be related to a particular skill performance, which is similar to a particular prototype.

The results provide further evidence that specific biomechanical properties, for example, an open hip angle, were reflected, not only in a single gymnastics movement, but also in the different gymnastics movements (Mack et al., in press). For the training process, this might imply that it would be useful to focus on the biomechanical property that should be improved. For instance, one training session could consist of practicing different skills with the same focus: back handspring and (backward and/or forward)somersault, stretched, focusing on the hip angle.

Another promising approach for diagnostics in the training process is neural networks, to determine, in real-time, the potential performance score that could be achieved, but also to identify the movement variables which should be changed to ensure a better performance (Mack et al., 2019). This could improve the training process by providing a tool for coaches and athletes, by supporting the diagnostic process and giving specific instructions regarding the movement components that should be changed.

Because no relationship could be found between the cluster solution of the kinematics pattern of the gymnastics skills and the cluster solution of their perceived movement quality (Mack et al., submitted), it is worth mentioning for the training process, that there are two explanations which should be kept in mind. One of them is that the evaluation of different movement characteristics has a different weight in the overall quality score. For example, a small deviation of the knee angle of the optimal value leads to a higher quality deduction than the same deviation of the hip angle from the optimal value. The other is that there are other performance-determined variables besides kinematics.

Judging Process

In his research, Sands (2011) mentioned that the biggest problem of evaluation is reliability and validity. The judgment score does not always reflect the real performance according to the criteria defined in the Code of Points, and its change from one evaluator to the next and from situation to another. This statement is in line with research showing that the evaluation process is influenced by variables, such as the judges' mental capacities, personal experiences, expertise, the point of view, the order of presentation and the subjectivity of the judging category (Ansorge et al., 1978; Heinen et al., 2013; Heiniger & Mercier, 2018; Pizzera, 2012; Pizzera & Raab, 2012; Plessner & Haar, 2006; Plessner & Schallis, 2005; Ste-Marie, 1999, Ste-Marie et al., 2001).

Results from the studies presented here can also be associated with those findings, for example, the unrelated kinematic pattern of presented skills with the perceived movement quality (Mack et al., submitted) or the importance of the last skill in the judgment process (Mack, 2019). Therefore, in judging courses, those potential influencing variables should be addressed, and the participants should be sensitized. Another idea for dealing with those issues was investigated by Pajek, Forbes, Pajek, Leskošek, & Čuk (2011). The program, B Jury Judging Real-Time System (RTJS) aimed to improve the objectivity of the evaluation. Jury B Execution Deductions are entered during the performance, and cannot be changed; judges must act quickly and precisely each time they see an error.

The use of modern technologies to improve the judgment process is getting more and more attention in gymnastics judging, not only because it is a tool with which judging can be done, but also because it replaces and supports judgments, based on a mathematical algorithm. One example is the judging support system based on movement data, which the International Gymnastics Federation wants to implement (The International Gymnastics Federation & Fujitsu Limited, 2017). Because the results suggest that it is potentially possible to predict judgment scores based on kinematic data, this is a promising new approach to improving judging in competitions. The use of neural networks in particular, might be appropriate, because it is not limited to a certain amount of data. It gets increasingly precise the more data it receives, and it can deal with a different weighting of different kinematics variables in the judgment.

Here, it should be mentioned that two factors should always be kept in mind when dealing with computer-based methodologies in the evaluation of performance. First, the psychological aspects. It could be assumed that being judged by a computer or a human being leads to differences regarding the execution of the skill. Each performance, especially those that are accompanied by music, distinguishes itself not only through a specific movement pattern, but also through emotional expression. It is questionable whether a mathematical algorithm is already able to capture this. Second, there is always the question of the error-rate in computer-based technologies. Yes, they are less influenced by individual psychological biases. However, on the other side, there are still a lot of different aspects (regarding capturing of motion, algorithms, etc.) that could lead to errors.

Chapter 11

Conclusion

In artistic gymnastics, the perceived performance quality plays an integral role for judges, coaches, athletes, and spectators. Thereby, one central question arises: How is the quantifiable execution of the movement related to the evaluation of the quality of the performance? The performance-determining guidelines (that describe biomechanical conditions) are thereby defined in the Code of Points, but the relationship is nevertheless not trivial, and it is influenced through different psychological variables. The aim of this thesis was therefore to kinematically analyze selected prototypical gymnastics skills, in order to investigate, in a second step, the underlying perceptual-cognitive mechanism when perceiving and evaluating those skills. In particular, the relationships between kinematics, gaze pattern, and the estimated quality of performance should be focused upon. In order to achieve this aim, five studies were conducted, each of which focused on different issues regarding action and perception-cognition. Theoretically and methodologically, all studies were based on the holistic nature of action and perception-action. Methods such as cluster-analysis (working with the relative similarity and proximity of variables), neural networks (a non-linear function computes input and output variables) as well as process-tracing methods such as eye-tracking, were used to this end.

The results provide first insights into how gymnastics performances and the subsequent judging scores are related, by focusing on the kinematics of the motion. The results should be handled with caution if they are used to draw practical conclusions, due to certain limitations in respect of the studies, such as the unvaried level of the movement data set, the confounding of the manipulation of expertise, and the operationalization of the judgment. But nevertheless, it is an initial approach, and it delivers the first promising results in a highly current and relevant domain, for two reasons: first, because the skills practiced in gymnastics are becoming faster and more complex, which seems to be a challenge for the limited human cognitive-visual system. In addition to the increasingly complex movements, the level of performance (especially on the high-elite level) is becoming more and more similar. Secondly, this challenge, and also the fast development of technologies, leads to the occurrence of computer-based judging systems. For instance, the International Gymnastics Federation is dealing with this issue by intensifying their efforts to implement a complex judging support system, based on movement data, in order to support the demand for the objective and potentially unbiased judgment of a gymnastics performance (The International Gymnastics Federation Fujitsu Limited, 2017).

We are now in an era in which technologies seem to develop at an almost frightening pace, and their influence on society and human life is unforeseeable. On the one hand, this entails great opportunities, but on the other hand, it also brings with it a lot of struggles and unpredictable difficulties. This is, of course, not only the case in gymnastics, but also in a lot of other domains (for instance, in learning analytics). From my point of view, and that is also an issue about which I feel very strongly, neither is it reasonable to make decisions exhaustively on the basis of those technologies. Gymnastics is not only about kinematics, and success is not about test-results at school. I would further recommend that those technologies not be strictly rejected. They could improve the objectivity of judgments and decisions in sports (but also in a lot of other domains, such as school, working life, and judicial decisions) and could lead to a more individual adjustment of training and learning processes.

References

Abadi, M. et al. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467.*

Abernethy B. (1990). Anticipation in squash: Differences in advance cue utilization between expert and novice players. *Journal of Sports Sciences, 8*, 17-34.

Abernethy, B., Gill, D. P., Parks, S. L., & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception, 30*(2), 233-252.

Abernethy, B., & Zawi, K. (2007). Pickup of essential kinematics underpins expert perception of movement patterns. *Journal of Motor Behavior, 39*(5), 353-367.

Abernethy, B., Zawi, K., & Jackson, R. C. (2008). Expertise and attunement to kinematic constraints. *Perception, 37*(5), 931-948.

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

Ansorge, C.J., Scheer, J.K., Laub, J., & Howard, J. (1978). Bias in judging women's gymnastics induced by expectations of within-team order. *Research Quarterly, 49*, 399-405.

Arkaev, L., & Suchilin, N. G. (2004). *Gymnastics: How to Create Champions.* Oxford, UK: Meyer Meyer Verlag.

Baker, F. B. (1974). Stability of two hierarchical grouping techniques case 1: Sensitivity to data errors. *Journal of the American Statistical Association, 69*(346), 440-445.

Bauer, H. U., & Schöllhorn, W. (1997). Self-organizing maps for the analysis of complex movement patterns. *Neural Processing Letters, 5*(3), 193-199.

Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics, 23*(2), 145-152.

Bard, C., Fleury, M., Carriere, L., & Halle, M. (1980). Analysis of gymnastics judges' visual search. *Research Quarterly for Exercise and Sport, 51*,267-273.

Bartlett, R. (1999). *Sports Biomechanics: Reducing Injury and Improving Performance.* London, UK: E & FN Spon.

Barton, G., Lees, A., Lisboa, P., & Attfield, S. (2006). Visualisation of gait data with Kohonen self-organising neural maps. *Gait & Posture, 24*(1), 46-53.

Behnke, R. S. (2001). *Kinetic Anatomy.* Champaign, IL: Human Kinetics.

Bernstein, N. (1967). *Coordination and regulation of movement.* Oxford, UK: Pergamon Press.

Bertenthal, B. I., Proffitt, D. R., Spetner, N. B., & Thomas, M. A. (1985). The development of infant sensitivity to biomechanical motions. *Child development, 56*(3), 531-543.

Beyer, E. (1992). *Wörterbuch der Sportwissenschaft.* Schorndorf, D: Hofmann.

Bläsing, B., & Schack, T. (2012). Mental representation of spatial movement parameters in dance. *Spatial Cognition & Computation,12*(2-3), 111-132.

Bläsing, B., Tenenbaum, G., & Schack, T. (2009). Cognitive structures of complex movements in dance. *Psychology of Sport and Exercise, 10*, 350-360.

Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology, 58*, 47-73.

Bradshaw, E., Hume, P., Calton, M., & Aisbett, B. (2010). Reliability and variability of day-to-day vault training measures in artistic gymnastics. *Sport Biomechanics, 9*, 79-97.

Bradshaw, E., & Sparrow, W. A. (2001). The approach, vaulting performance, and judge's score in women's artistic gymnastics. In J. R. Blackwell & R. H. Sanders (Eds.), *XIX International Symposium of Biomechanics in Sports* (pp. 311-314). San Francisco, CA: University of San Francisco.

Braun, S. M., Beurskens, A. J., Schack, T., Marcellis, R. G., Oti, K. C., Schols, J. M., & Wade, D. T. (2007). Is it possible to use the Structural Dimension Analysis of Motor Memory (SDA-M) to investigate representations of motor actions in stroke patients? *Clinical Rehabilitation, 21*(9), 822-832.

Brüggemann, G.-P. (1994). Biomechanics of gymnastic techniques. *Sport Science Review, 3*(2), 79-120.

Bryan M., Heyer G., Philipp N., Rehak M., & Wiedemann P. (2020) Convolutional attention on images for locating macular edema. In: Y. Zheng, B. Williams, & Chen K. (Eds.), *Medical Image Understanding and Analysis. MIUA 2019. Communications in Computer and Information Science, vol 1065.* Cham: Springer.

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition, 2*(1).

Button, C., MacLeod, M., Sanders, R., & Coleman, S. (2003). Examining movement variability in the basketball free-throw action at different skill levels. *Research Quarterly for Exercise and Sport, 74*, 257-269.

Calvo-Merino, B., Ehrenberg, S., Leung, D., & Haggard, P. (2010). Experts see it all: configural effects in action observation. *Psychological Research, 74*(4), 400-406.

Cañal-Bruland, R., & Williams, A. M. (2010). Movement recognition and prediction of movement effects in biological motion perception. *Experimental Psychology, 57*(4), 320-326.

Carr, G. (1997). *Mechanics of Sport.* Champaign, IL: Human Kinetics.

Cicchetti, D. V. (1994). Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290.

Chollet, F. (2015). Keras. Retrived from https://keras.io

Chow, J. W., & Knudson, D. V. (2011). Use of deterministic models in sports and exercise biomechanics research. *Sports Biomechanics, 10*(3), 219-233.

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555.*

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. (2018). A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports Medicine - Open, 4*, 24.

Creswell, J. Tashakkori, A. (2007). Differing perspectives on mixed methods research. *Journal of Mixed Methods Research, 1*(4), 303-308.

Čuk, I., (2015). Can audience replace execution judges in male gymnastics? *Science of Gymnastics Journal, 7*(1), 61-68.

Čuk, I., & Karacàsony, I. (2004). *Vault. Methods, ideas, curiosities, history.* Slovenia, SI: ŠTD Sangvinčki.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society, 9*, 353-356.

Cutting, J. E., Proffitt, D. R., & Kozlowski, L. T. (1978). A biomechanical invariant for gait perception. *Journal of Experimental Psychology: Human Perception and Performance, 4*(3), 357-372.

Daffertshofer, A., Lamoth, C. J., Meijer, O. G., & Beek, P. J. (2004). PCA in studying coordination and variability: a tutorial. *Clinical Biomechanics, 19*(4), 415-428.

Davids, K., Button, C., & Bennett, S. J. (2008). *Dynamics of skill acquisition: A constraints-led approach*, Champaign, IL: Human Kinetics.

Davids, K., Hristovski, R., Araújo, D., Serre, N. B., Button, C., & Passos, P. (2014). *Complex systems in sport.* New York, NY: Routledge.

de Bruin, W. B. (2006). Save the last dance II: Unwanted serial position effects in figure skating judgments. *Acta Psychologica, 123*(3), 299-311.

del Campo, V. L., & Gracia, I. E. (2018). Exploring visual patterns and judgments predicated on role specificity: Case studies of expertise in gymnastics. *Current Psychology, 37*(4), 934-941.

Díaz-Pereira, M. P., Gomez-Conde, I., Escalona, M., & Olivieri, D. N. (2014). Automatic recognition and scoring of olympic rhythmic gymnastic movements. *Human Movement Science, 34*, 63-80.

Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception, 22*(1), 15-22.

Dosseville, F., Laborde, S., & Garncarzyk, C. (2014). Current research in sports officiating and decision-making. In C. Mohiyeddini (Ed.), *Contemporary Topics and Trends in the Psychology of Sports* (pp. 13-38). New York: Nova Publishers.

DTB. *Handbuch des Deutschen Turner-Bundes, Teil 1, Aufgabenbuch, Broschüre 1 - Gerätturnen weiblich [Handbook of the German Gymnastics Federation, Part 1, compulsory exercises, brochure 1, female artistic gymnastics].* 3rd ed. Frankfurt/M., Germany: Deutscher Turner-Bund Service GmbH. 2001.

Elliott, B.C. (1982). Tennis: the influence of grip tightness on reaction impulse and rebound velocity. *Medicine and Science in Sports and Exercise, 14*, 348-352.

Enoka, R. M. (2002). *Neuromechanics of human movement (3rd ed.).* Champaign, IL: Human Kinetics.

Endres, F., Hess, J., & Burgard, W. (2012) Graph-based action models for human motion classification. *Proceedings of ROBOTIK 2012, 7th German Conference on Robotics.* (pp. 1-6). Munich, GE: VDE.

Everitt, B. S., & Dunn, G. (2001). *Applied multivariate data analysis (2nd ed.).* London, UK: Arnold.

Farana, R., Uchytil, J., Zahradník, D., & Jandacka, D. (2015). The "Akopian" vault performed by elite male gymnasts: which biomechanical variables are related to a judge's score. *Acta Gymnica, 45*(1), 33-40.

Farana, R., & Vaverka, F. (2012). The effect of biomechanical variables on the assessment of vaulting in top-level artistic female gymnasts in World Cup competitions. *Acta Gymnica, 42*(2), 49-57

Federolf, P., Tecante, K., & Benno. (2012). A holistic approach to study the temporal variability in gait. *Journal of Biomechanics, 45*(8), 1127-1132.

Fédèration Internationale de Gymnastique [FIG] (2017). 2017 - 2020 Code of Points. Women's artistic gymnastics. Retrieved from https://www.fig-gymnastics.com/site /rules/rules.php

Federolf, P ., Reid, R., Gilgien, M., Haugen, P., & Smith, G. (2014). The application of principal component analysis to quantify technique in sports. *Scandinavian Journal of Medicine & Science in Sports, 24*(3), 491-499.

Ferger, K., & Hackbarth, M. (2017). New way of determining horizontal displacement in competitive trampolining. *Science of Gymnastics Journal, 9*, 303-310.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R.* London: Sage publications.

Fitzpatrick, P., Schmidt, R. C., & Lockman, J. J. (1996). Dynamical patterns in the development of clapping. *Child Development, 67*(6), 2691-2708.

Fransella, F., Bell, R., & Bannister, D. (2004). *A manual for repertory grid technique (2nd ed.).* Chichster, UK: John Wiley & Sons.

Galili, T. (2015). dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering. Bioinformatics.

Galley, N., Betz, D., & Biniossek, C. (2015). Fixation durations: Why are they so highly variable? In T. Heinen (Ed.), *Advances in visual perception research* (pp. 83-106). New York, NY: Nova Science Publishers Inc.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review, 23*(4), 523-552.

Giblin, G., Farrow, M.R., Ball, K., & Abernethey, B. (2015). Perceiving movement patterns: Implications for skill evaluation, correction and development. *RICYDE. Revisit International de Ciencias Del Deporte, 39*(11), 5-17.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*(1), 1-19.

Glöckner, A., Heinen, T., Johnson, J. G., & Raab, M. (2012). Network approaches for expert decisions in sports. *Human Movement Science, 31*(2), 318-333.

GymnasticsResultsCom. (December, 2019). Retrieved from www.gymnasticsresults.com

Handford, C. (2006). Serving up variability and stability. In K. Davids, S. Bennett, & K. Newell (Eds.), *Movement system variability* (pp. 73-83). Champaign, IL: Human Kinetics.

Hars, M., & Calmels, C. (2007). Observation of elite gymnastic performance: Processes and perceived functions of observation. *Psychology of Sport and Exercise, 8*(3), 337-354.

Hausken, K. (2017). Exhaustive classification and review of techniques and research program for techniques for Skate Skiing, Classical Skiing, and Ski Mountaineering. *The Open Sports Sciences Journal, 10*(1). 160-178.

Hauw, D., Renault, G., & Durand, M. (2008). How do aerial freestyler skiers land on their feet? A situated analysis of athletes' activity related to new forms of acrobatic performance. *Journal of Science and Medicine in Sport, 11*(5), 481-486.

Hay, J.G., & Reid, J.G. (1988). *Anatomy, Mechanics and Human Motion.* Englewood Cliffs, NJ: Prentice Hall.

Heiniger, S., & Mercier, H. (2018). National Bias of International Gymnastics Judges during the 2013-2016 Olympic Cycle. *arXiv preprint arXiv:1807.10033.*

Heinen, T., Mandry, S., Vinken, P., & Nicolaus, M. (2013). Motor skill acquisition influences learners' visual perception in gymnastics. *Science of Gymnastics Journal, 5*, 19-28.

Heinen, T., Vinken, P. M., & Velentzas, K. (2012). Judging performance in gymnastics: a matter of motor or visual experience? *Science of Gymnastics Journal, 4*(1), 63-72.

Higgins, J. R. (1977). *Human movement: an integrated approach.* St. Louis, MO: Mosby.

Hiley, M. J., Wangler, R., & Predescu, G. (2009). Optimization of the felge on parallel bars. *Sports Biomechanics, 8*(1), 39- 51.

Hochmuth, G., & Marhold, G. (1978). Zur Weiterentwicklung der biomechanischen Prinzipien [Further development of biomechniacal principles]. *Theorie und Praxis der Körperkultur, 27*, 202-208.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation, 9*(8), 1735-1780.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*(5), 849-878.

Horn, R. R., Williams, A. M., & Scott, M. A. (2002). Learning from demonstrations: the role of visual search during observational learning from video and point-light models. *Journal of Sports Sciences, 20*(3), 253-269.

Hudson, J. L. (1985). Prediction of basketball skill using biomechanical variables. *Research Quarterly, 56*, 115-121.

Huys, R., Smeeton, N. J., Hodges, N. J., Beek, P., & Williams, A. M. (2008). The dynamical information underlying anticipation skill in tennis. *Perception and Psychophysics, 18*, 1217-1234.

Irwin, G., Manning, M., & Kerwin, D. G. (2011). Kinematics and angular momentum contributions to the toe-on Tkachev on uneven bars in female gymnastics. *Biomechanics in Sports, 29*, 11 (suppl. 2).

Jaitner, T., Mendoza, L., & Schöllhorn, W. I. (2001). Analysis of the long jump technique in the transition from approach to takeoff based on time-continuous kinematic data. *European Journal of Sport Science, 1*(5), 1-12.

Jeraj, D. (2016). Error perception in gymnastics: two consecutive interventions. *Science of Gymnastics Journal, 8*, 239-253.

Jeraj, D., Hennig, L., & Heinen, T. (2015). The error-correction process - a heuristic concept. In T. Heinen (ed.), *Advances in Visual Perception Research (pp. 193-202).* New York, NY: Nova Science Publishers Inc.

Janelle, C.M., & Hillman, C.H. (2003). Expert performance in sport: Current perspective and critical issues. In J.L. Starkes & K.A. Ericsson (Eds.), *Expert Performance in sports: Advances in research on sport expertise.* Champaign, IL: Human Kinetics.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14*(2), 201-211.

Johansson, G. (1975). Visual motion perception. *Scientific American, 232*(6), 76-89.

Johnson-Laird, P. N. (1989). Mental models. In M. I. Posner (Ed.), *Foundations of Cognitive Science* (pp. 469-499). Cambridge, MA: MIT Press.

Karàcsony, I., & Čuk, I. (2005). *Floor exercises - Methods, ideas, curiosities, history.* Ljubljana, SI: STD Sangvincki.

Kelso, J. S. (1997). *Dynamic patterns: the self-organization of brain and behavior.* Cambridge, MA: MIT press.

Kermarrec, G., Todorovich, J., & Fleming, D. (2004). An investigation of the self-regulation components students employ in the physical education setting. *Journal of Teaching in Physical Education, 23*, 123-142.

Kerwin, D. G., & Irwin, G. (2010). Musculoskeletal work preceding the outward and inward Tkachev on uneven bars in artistic gymnastics. *Sports Biomechanics, 9*(1), 16-28.

Kim, T., Park, H., & Schack, T. (2019). A functional link between mental representation in long-term memory and cognitive performance in working memory. *Advances in cognitive Psychology, 15*(1). 11-20.

King, M. A., & Yeadon, M. R. (2004). Maximising somersault rotation in tumbling. *Journal of Biomechanics, 37*(4), 471-477.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine, 15*(2), 155-163.

Land, W., Volchenkov, D., Bläsing, B. E., & Schack, T. (2013). From action representation to action execution: Exploring the links between cognitive and biomechanical levels of motor control. *Frontiers in Computational Neuroscience, 7*, 127.

Lander, H.-J. (1991). Ein methodischer Ansatz zur Ermittlung der Struktur und der Dimensionierung einer intern repräsentierten Begriffsmenge. [A methodological approach to determining the structure and dimensioning of internally represented concepts]. *Zeitschrift für Psychologie, 199*(2), 167-176.

Lander, H. J., & Lange, K. (1996). Untersuchung zur Struktur- und Dimensionsanalyse begrifflich repräsentierten Wissens. *Zeitschrift für Psychologie, 204*, 55-74.

Latash, M. L., Scholz, J. P., & Schöner, G. (2002). Motor control strategies revealed in the structure of motor variability. *Exercise and Sport Sciences Reviews, 30*(1), 26-31.

Lees, A. (2002). Technique analysis in sports: a critical review. *Journal of Sports Science, 20*, 813-828.

Leskošek, B., Čuk, I., Karácsony, I., Pajek, J., & Bučar, M. (2010). Reliability and validity of judging in men's artistic gymnastics at the 2009 university games. *Science of Gymnastics Journal, 2*(1), 25-34.

Li, L., van den Bogert, E. C., Caldwell, G. E., van Emmerik, R. E., & Hamill, J. (1999). Coordination patterns of walking and running at similar speed and stride frequency. *Human Movement Science, 18*(1), 67-85.

Leirich, J., & Rieling, K. (1969). Zur strukturelle Anordnung der übungen des Gerätturnens: Die Sprungbewegungen [The structure of the exercises in gymnastics: Jumps]. *Theorie und Praxis der Körperkultur, 18*, 139-147.

Leskošek, B., Čuk, I., Karácsony, I., Pajek, J., & Bučar, M. (2010). Reliability and validity of judging in men's artistic gymnastics at the 2009 university games. *Science of Gymnastics Journal, 2*(1), 25-34.

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 210.

MacMahon, C., Helsen, W. F., Starkes, J. L., & Weston, M. (2007). Decision-making skills and deliberate practice in elite association football referees. *Journal of Sports Sciences, 25*(1), 65-78.

MacMahon, C., Mascarenhas, D., Plessner, H., Pizzera, A., Oudejans, R. & Raab, M. (2015). *Sports Officials and Officiating - Science and Practice.* Abingdon, UK: Routledge.

Mack, M. (2020). Exploring cognitive and perceptual judgment processes in gymnastics by using essential kinematic information. *Advances in Cognitive Psychology, 16*(1), 34-44.

Mack, M., Bryan, M., Heyer, G., & Heinen, T. (2019). Modeling judges' scores in artistic gymnastics. *The Open Sports Sciences Journal, 12*, 1-9.

Mack, M., Federbusch, S., Ferber, M., & Heinen, T. (in press). Movement prototypes and their relationship in the performance of a gymnastics floor routine. *Journal of Human Sport and Exercise, 15.*

Mack, M., Hennig, L., & Heinen, T. (2018). Movement prototypes in the performance of the handspring on vault. *Science of Gymnastics Journal, 10*(2), 245-257.

Mack, M., Schmidt, M., & Heinen, T. (submitted). Mental representation of movement quality of complex skills in gymnastics.

Mann, D. L., Causer, J., Nakamoto, H., & Runswick, O. R. (2019). Visual search behaviors in expert perceptual judgements. In A. M. Williams, & R. C. Jackson (Eds.). *Anticipation and decision making in sport* (pp. 59-78). Abingdon, Oxon: Routledge.

Mann, D. T., Williams, A. M., Ward, P., & Janelle, C. M.(2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology, 29*(4), 457-478.

Massidda, M., & Calò, C. M. (2012). Performance scores and standings during the 43rd Artistic Gymnastics World Championships, 2011. *Journal of Sports Sciences, 30*(13), 1415-1420.

McGinnis, P. M. (2005). *Biomechanics of sport and exercise (2nd ed.).* Champaign, IL: Human Kinetics.

Mercier, H., & Heiniger, S. (2018). Judging the judges: Evaluating the performance of international gymnastics judges. *arXiv preprint arXiv:1807.10021.*

Miller, S. A. (1998). The kinematics of inaccuracy in basketball shooting. In H. J. Riehle and M. M. Vieten (eds), *Proceedings of XVI International Symposium on Biomechanics in Sports* (pp. 188-191). Konstanz; D: Universitätsverlag Konstanz.

Miller, S. A. (2002). Variability in basketball shooting: Practical implications. In Y. Hong (ed.), *International Research in Sports Biomechanics* (pp. 27?34). London, UK: Routledge.

Miller, S. A., & Bartlett, R. M. (1996). The relationship between basketball shooting kinematics, distance and playing position, *Journal of Sports Sciences, 14*, 243-253.

Moreno, F. J., Reina, R., Luis, V., & Sabido, R. (2002). Visual search strategies in experienced and inexperienced gymnastic coaches. *Perceptual and Motor Skills, 95*,901-902.

Morriss, C., Bartlett, R., & Navarro, E. (2001). The function of blocking in elite javelin throws: a re-evaluation. *Journal of human movement studies, 41*(3), 175-190.

Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement ward's criterion? *Journal of Classification, 31*(3), 274-295.

Naundorf, F., Brehmer, S., Knoll, K., Bronst, A. & Wagner, R. (2008). Development of velocity for vault runs in artistic gymnastics for last decade. In Y.H. Kwon, J. Shim, J.K. Shim, and I.S. Shin (eds.), *XXVI International Symposium of Biomechanics in Sports.* Seoul: Korea.

Newell, K. M. (1985). Coordination, control and skill. In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds.), *Differing Perspectives in Motor Learning, Memory, and Control* (pp. 295-317). Amsterdam, NL: Elsevier Science.

Newell, K. M., & Jordan, K. (2007). Task constraints and movement organization: A common language. In W. E. Davis, and G. D. Broadhead (Eds.), *Ecological Task Analysis and Movement* (pp. 5-23). Champaign, IL: Human Kinetics.

Newell, K. M., Deutsch, K. M., Sosnoff, J., & Mayer-Kress, G. (2006). Variability in motor output as noise: A default and erroneous proposition. In K. Davids, S. Bennett, & K. M. Newell (Eds.), *Variability in the movement System: A multidisciplinary perspective* (pp. 3-23). Champaign, IL: Human Kinetics.

Omorczyk J., Nosiadek L., Ambroży T., & Nosiadek A. (2015). High-frequency video capture and a computer program with frame-by-frame angle determination functionality as tools that support judging in artistic gymnastics. *Acta of Bioengineering Biomechanics, 17*(3), 85-93.

Oullier, O., Martin, L., Stoffregen, T., Bootsma, R., & Bardy, B. (2006): Variability in Postural Coordination Dynamics. In Davids, K., Bennett, S. und Newell, K. (Eds.), *Movement System Variability* (pp. 25-47). Champaign, IL: Human Kinetics.

Pajek, M. B., Forbes, W., Pajek, J., Leskošek, B., & Čuk, I. (2011). Reliability of real time judging system. *Science of Gymnastics Journal, 3*(2), 47-54.

Pajek, M. B., Čuk, I., Pajek, J., Kovač, M., & Leskošek, B. (2013). Is the quality of judging in women artistic gymnastics equivalent at major competitions of different levels? *Journal of Human Kinetics, 37*(1), 173-181.

Penitente, G., Merni, F., & Fantozzi, S. (2009). Onboard and pre-flight mechanical model of Yurchenko one twist on vault: Implications for performance. In A. J. Harrison, R. Anderson, & I. Kenny (Eds.), *XXVII International Symposium of Biomechanics in Sports* (pp. 636-639). Limerick, IRL: International Society of Biomechanics in Sports.

Penrose, T. & Blanksby, B. (1976). Film analysis: Two methods of basketball jump shooting techniques by two groups of different ability levels. *The Australian Journal for Health, Physical Education and Recreation, 71*, 14-21.

Pizzera, A. (2012). Gymnastic judges benefit from their own motor experience as gymnasts. *Research Quarterly for Exercise and Sport, 83*(4), 603-607.

Pizzera, A., Möller, C., & Plessner, H. (2018). Gaze behavior of gymnastics judges: Where do experienced judges and gymnasts look while judging? *Research quarterly for exercise and sport, 89*(1), 112-119.

Pizzera, A., & Raab, M. (2012). Perceptual judgments of sports officials are influenced by their motor and visual experience. *Journal of Applied Sport Psychology, 24*(1), 59-72.

Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognitive perspective. *Psychology of Sport and Exercise, 7*(6), 555-575.

Plessner, H., & Schallies, E. (2005). Judging the cross on rings: A matter of achieving shape constancy. *Applied Cognitive Psychology, 19*(9), 1145-1156.

Poizat, G., Bourbousson, J., Saury, J., & Sève, C. (2012). Understanding team coordination in doubles table tennis: Joint analysis of first-and third-person data. *Psychology of Sport and Exercise, 13*(5), 630-639.

Prassas, S., Kwon, Y. H., & Sands, W. A. (2006). Biomechanical research in artistic gymnastics: A review. *Sports Biomechanics, 5*(2), 261-291.

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

Prinz, W. (1990). A common coding approach to perception and action. In *Relationships between perception and action* (pp. 167-201). Springer, Berlin, Heidelberg.

R Core Team (2017). *R: A language and environment for statistical computing. R foundation for statistical computing*, Vienna, Austria. Retrived from https://www.r-project.org/

Rein, R., Button, C., Davids, K., & Summers, J. (2010). Cluster analysis of movement patterns in multiarticular actions: A tutorial. *Motor Control, 14*(2), 211-239.

Ren, L., Patrick, A., Efros, A. A., Hodgins, J. K., & Rehg, J. M. (2005). A data-driven approach to quantifying natural human motion. In *ACM Transactions on Graphics (TOG),24* (3), 1090-1097.

Robins, M., Wheat, J.S., Irwin, G., & Bartlett, R. (2006). The effect of shooting distance on movement variability in basketball. *Journal of Human Movement Studies, 20*, 218-238.

Romack, J. L. (1995). Information in visual event perception and its use in observational learning. In B. G. Bardy, R. J. Bootsma & Y. Guiard (Eds.), *Studies in perception and action II* (pp. 289-292). Hillsdale, NJ: Erlbaum.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General, 112*, 585-615.

Salamah S., Zhang L., & Brunnett G. (2015) Hierarchical method for segmentation by classification of motion capture data. In G. Brunnett, S. Coquillart, R. van Liere, G. Welch, & L. Váša (Eds.), *Virtual Realities. Lecture Notes in Computer Science, vol 8844.* Cham, CH: Springer.

Sands, W. A. (2011). Angular kinetics applied to gymnastics. In M. Jemni (Ed.), *The Science of Gymnastics* (pp. 94-104). London, UK: Routledge.

Schack, T. (2003). The relationship between motor representation and biomechanical parameters in complex movements: towards an integrative perspective of movement science. *European Journal of Sport Science, 3*(2), 1-13.

Schack, T. (2004). The cognitive architecture of complex movement. *International Journal of Sport and Exercise Psychology, 2*(4), 403-438.

Schack, T. (2010). *Die kognitive Architektur menschlicher Bewegungen: Innovative Zugänge für Psychologie, Sportwissenschaft und Robotik.* Oxford, UK: Meyer & Meyer Verlag.

Schack, T., & Mechsner, F. (2006). Representation of motor skills in human long-term memory. *Neuroscience letters, 391*(3), 77-81.

Schack, T., & Ritter, H. (2009). The cognitive nature of action - functional links between cognitive psychology, movement science, and robotics. *Progress in Brain Research, 174*, 231-250.

Schöllhorn, W. I., & Bauer, H.U. (1998). Identifying individual movement styles in high performance sports by means of self-organizing Kohonen maps. In H. J. Riehle & M. Vieten (Eds.), *Proceedings of the XVI ISBS 1998* (pp. 574-577). Konstanz, GE: University Press.

Schöllhorn, W.I., Mayer-Kress, G., Newell, K.M., & Michelbrink, M. (2009). Time scales of adaptive behavior and motor learning in the presence of stochastic perturbations. *Human Movement Science, 28*, 319-333.

Schöllhorn, W., Chow, J. Y., Glazier, P., & Button, C. (2014). Self-organising maps and cluster analysis in elite and sub-elite athletic performance. In K. Davids, R. Hristovski, D. Araújo, N. Balagué Serre, C. Button, & P. Passos (Eds.), *Complex Systems in Sport* (pp. 145-159). London: Routledge.

Scully, D., & Carnegie, E. (1998). Observational learning in motor skill acquisition: A look at demonstrations. *The Irish Journal of Psychology, 19*(4), 472-485.

Simi Reality Motion Systems GmbH. *Simi Motion®*. Unterschleißheim, Germany.

Ste-Marie, D. M. (1999). Expert-novice differences in gymnastic judging: an information-processing perspective. *Applied Cognitive Psychology, 13*(3), 269-281.

Ste-Marie, D. M. (2000). Expertise in women's gymnastic judging: An observational approach. *Perceptual and Motor Skills, 90*, 543-546.

Ste-Marie, D. M., Valiquette, S. M., & Taylor, G. (2001). Memory-influenced biases in gymnastic judging occur across different prior processing conditions. *Research Quarterly for Exercise and Sport, 72*(4), 420-426.

Stergiou, N. (2004). *Innovative analysis of human movement.* Champaign, IL: Human Kinetics.

Ste-Marie, D. M. (2003). Expertise in sport judges and referees: Circumventing information-processing limitations. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports. Advances in research on sport expertise* (pp. 169-190). Champaign, IL: Human Kinetics.

Stevenage, S. V., Nixon, M. S., & Vince, K. (1999). Visual analysis of gait as a cue to identity. *Applied Cognitive Psychology, 13*, 513-526.

Stöckel, T., Hughes, C. M. L., & Schack, T. (2012). Representation of grasp postures and anticipatory motor planning in children. *Psychological Research, 76*, 768-776.

Takei, Y. (1990). Technique used by elite women gymnasts performing a handspring vault at the 1987 Pan American Games. *International Journal of Sport Biomechanics, 6*, 29-55.

Takei, Y. (1992). Blocking and post flight technique of male gymnasts performing the compulsory vault at the 1988 Olympics. *International Journal of Applied Biomechanics, 7*, 87-110.

Takei, Y. (1998). Three-dimensional analysis of handspring with full turn vault: Deterministic model, coaches' beliefs, and judges' scores. *Journal of Applied Biomechanics, 14*, 190-210.

Takei, Y. (2007). The Roche vault performed by elite gymnasts: Somersaulting technique, deterministic model, and judges' scores. *Journal of Applied Biomechanics, 23*(1), 1-11.

Takei, Y., Blucker, E. P., Nohara, H., & Yamashita, N. (2000). The Hecht vault performed at the 1995 World Gymnastics Championships: deterministic model and judges' scores. *Journal of Sports Sciences, 18*(11), 849-863.

Takei, Y., & Dunn, J. H. (1996). A comparison of techniques used by elite gymnasts in performing the basket-to-handstand mount. *Journal of Sports Sciences, 14*(3), 269-279.

Takei, Y., & Kim, E. J. (1990). Techniques used in performing the handspring and salto forward tucked vault at the 1988 Olympic Games. *Journal of Applied Biomechanics, 6*(2), 111-138.

The International Gymnastics Dederation & Fujitsu Limited (2017). The International Gymnastics Federation and Fujitsu to collaborate on building a judging support system for artistic gymnastics competitions. Fujitsu Press releases. Retrieved from www.fujitsu.com /global/about/resources/news/press-releases/2017/1007-01.html

Thompson, J. M. T., & Stewart, H. B. (2002). *Nonlinear dynamics and chaos.* New York, NY: John Wiley & Sons.

Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision, 2*(5), 371-387.

Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics, 67*(4), 667-675.

Turoff, F. (1991). *Artistic gymnastics. A comprehensive guide to performing and teaching skills for beginners and advanced beginners.* Dubuque, IA: Wm. C. Brown Publishers.

Van der Eb, J., Filius, M., Ruogoor, G., Van Niel, C., de Water, J., Coolen, B., & de Koning, H. (2012). Optimal velocity profiles for vault. In E. Bradshaw, A. Burnett, & P. Hume (Eds.), *XXX International Symposium of Biomechanics in Sports* (pp. 71-75). Melbourne, AUS: International Society of Biomechanics in Sports.

Veličković, S., Petković, D., & Petković, E. (2011). A case study about differences in characteristics of the run-up approach on the vault between top-class and middle-class gymnasts. *Science of Gymnastics Journal, 3*, 25-34.

Vickers, J. N. (1988). Knowledge structures of expert-novice gymnasts. *Human Movement Science, 7*(1), 47-72.

Vickers, J. N. (2007). *Perception, cognition, and decision training: The quiet eye in action.* Champaign, Il: Human Kinetics.

Vosgerau, G. (2009). *Mental representation and self-conscious-ness: From basic self-representation to self-related cognition.* Münster, GE: Mentis Verlag.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*(301), 236-244.

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

Warren, W. H. (2006). The dynamics of perception and action. *Psychological Review, 113*(2), 358-389.

Williams, A.M., Davids, K., & Williams, J.G. (1999). *Visual Perception and Action in Sport.* London, UK: E & FN Spon.

Williams, A. M., & Ericsson, K. A. (2005). Perceptual-cognitive expertise in sport: Some considerations when applying the expert performance approach. *Human Movement Science, 24*, 283-307.

Williams, A. M., Huys, R., Cañal-Bruland, R., & Hagemann, N. (2009). The dynamical information underpinning deception effects. *Human Movement Science, 28*, 362-370.

Williams, G. K. R., Irwin, G., Kerwin, D. G., & Newell, K. M. (2015). Kinematic changes during learning the longswing on high bar. *Sports Biomechanics, 11*, 20-33.

Williams, G. K. R., Irwin, G., Kerwin, D. G., & Newell, K. M. (2015). Biomechanical energetic analysis of technique during learning the longswing on the high bar. *Journal of Sports Sciences, 33*(13), 1376-1387.

Young, C., & Reinkensmeyer, D. J. (2014). Judging complex movement performances for excellence: a principal components analysis-based technique applied to competitive diving. *Human Movement Science, 36*, 107-122.

211UKRjump (2019, January 07). *Simone Biles (USA) Floor Team Finals 2016 Rio Olympic Games* [Video file]. Retrieved from https://www.youtube.com/watch?v=KwhH-rvuaYQ

Appendix

- Versicherung

Versicherung

Hiermit versichere ich, Melanie Mack, dass ich die vorliegende Arbeit, mit dem Titel "Performing and Observing Complex Skills in Gymnastics: An Investigation of Prototypical Movement Patterns and Perceptual-Cognitive Processes", ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts habe ich Unterstützungsleistungen von folgenden Personen erhalten: Maximilian Bryan, Sarah Federbusch, Melanie Ferber, Thomas Heinen, Linda Hennig, Gerhard Heyer, Marcel Schmidt

Weitere Personen waren an der geistigen Herstellung der Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prfüngsbehörde vorgelegt und ist auch noch nicht veröffentlicht worden.

Unterschrift: ______________________________

Datum: ______________________________