

Exploring Cognitive and Perceptual Judgment Processes in Gymnastics Using Essential Kinematics Information

Melanie Mack

Faculty of Sport Science, Leipzig University, Leipzig, Germany

ABSTRACT

The aim of the current study was to transfer the method of using isolated kinematics information combined with the examination of perceptual-cognitive processes to gymnastics judging and thereby investigating the informational underpinnings 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 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 analyzed for the whole floor routine as well as for the particular temporal-spatial 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 judgment accuracy, $\eta_p^2 = .196$, but not 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 judgment instruction on the gaze pattern as well as the importance of the last skill for the judgment score.

KEYWORDS

eye tracking
judgment
gymnastics
expertise
perception

INTRODUCTION

The judgment of gymnastics skills, often under conditions of limited time and restricted resources, seems to be a complex task, and 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 large, as well as crowded with information both relevant and 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 underpinnings 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 or stick-figure format that 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 together with artistic elements. Like other gymnastics disciplines, for example, synchronized swimming or figure skating, the performance depends

Corresponding author: Melanie Mack, Faculty of Sport Science, Leipzig University, Jahnallee 59, 04109 Leipzig, Germany. E-mail: melanie.mack.m@gmail.com

exclusively on the judges' evaluation, based on criteria defined in the code of points (Arkaev & Suchilin, 2004; Fédération Internationale de Gymnastique [FIG], 2017). This is a regular compendium providing the kinematics criteria by which the performances should be evaluated. Therefore, the kinematics of the performed skill has to be compared with the kinematics of an optimal execution model (Ste-Marie, 1999). The variability of gymnastics judgments and its underlying factors are described and investigated in a large body of literature to identify and better understand the variables in the evaluation process where bias could occur (Ansoerge et al., 1978; Heinen et al., 2013; Pizzera, 2012; Pizzera & Raab, 2012; Plessner & Haar, 2006; Plessner & Schallies, 2005; MacMahon & Plessner, 2008; 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., 2013; Prinz, 1997), raising the key question of how the quantifiable movement execution is related to the evaluation of performance quality. Mixed methods approaches, providing the opportunity to link the investigation of those different variables and thereby to better understand their relationship, are still rare. The term mixed methods thus refers to the method perspective of collecting, analyzing, and interpreting multiple types of quantitative and qualitative data (Creswell & Tashakkori, 2007).

In general, it is supposed that the process of perceiving and evaluating motion uses relative information imparted through the motion itself. More specifically, kinematic cues over time and space are integrated into the evaluation process (Johansson, 1973; Troje, 2002). There is also a body 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 using deterministic models (e.g., 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 of certain visual search strategies, the anticipation of extended movement cues, recognition of movement pattern, and the use of probabilistic information.

One method to investigate the influence of kinematics in the evaluation process is the use of point-light displays. It involves representing major joints of a moving person with dots. Research shows that not only particular actions (Dittrich, 1993), but also the gender of a walking person (Barclay et al., 1978; Cutting et al., 1978; Troje, 2002) as well as individual persons (Cutting & Kozlowski, 1977; Stevenage et al., 1999; Troje et al., 2005) can be identified through a point-light display. This method was also used in the sports context, where dynamic information was manipulated or occluded and presented to the participants. It was shown that information relevant for the evaluation could be extracted out of the point-light displays (Cañal-Bruland & Williams, 2010; Huys et al., 2008; Williams et al., 2009). Thus, the importance

of inferring postural cues and biological motion when attempting to anticipate an opponent's 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 information search. It can provide important insights to 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 pattern as athletes attempt to anticipate or judge skill performance in both laboratory and field settings (Mann et al., 2019). Several indices of gaze were found which seem to influence performance by indexing an individual's point of interest and information processing, for example, the location, the amount, and the duration of fixations. Those indices were used to operationalize certain information-processing assumptions. For example, a longer fixation duration seems to lead to a more detailed information processing (but not necessarily of the fixated visual cues) by allowing an extraction of more 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).

Many studies have shown that there are expertise differences concerning the judgment of performances in gymnastics, but also the encoding, processing, and retrieving of information (Hars & Calmels, 2007; Heinen et al., 2012; Ste-Marie et al., 2001). Expert judges are better in error detection (Bard et al., 1980) and perception of body angles (Plessner & Schallies, 2005) compared to novices. Also, the perception-action link (common coding theory, Prinz, 1990) states that not only visual but also bodily experiences influence perception and judgment and should be taken into account (Pizzera & Raab, 2012). For instance, motoric knowledge of a skill leads to a more accurate estimation of joint angles and detection of form errors than only 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 body than novices. (Abernethy et al., 2008). Evidence for a general strong link between perceptual expertise and the kinematics of the observed action comes from studies demonstrating expert-novice differences in the time course of information acquisition (Abernethy et al., 2001) and anticipatory performances (Abernethy & Zawi, 2007; Ward et al., 2002) by using point-light displays.

The investigation of gaze patterns showed that different gymnastics expertise leads to a different patterns as well as to different judgments. Participants focused on different information sources when making judgments on gymnastics performance based on their 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 concerning the underlying mechanism of perceptual skills in sports, Mann et al. (2007) showed that, in general, experts are better in 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, skill-based differences in gaze pattern were found with experts scanning the display more systematically and fixating on different areas than novices. However, the investigation of gaze patterns possess limitations because of peripheral perception. The point of fixation is not necessarily the point of attention. This leads to the need for 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., 2013; Prinz, 1997). Thus, the question arises of how the quantifiable movement execution is related to the evaluation of performance quality.

Currently, there is a growing interest for mixed methods in the field of sport science, providing the opportunity to link the investigation of those different variables and thereby better understand their relationship (Poizat et al., 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, 49 participants with different levels of gymnastics expertise were recruited for the study. The number of participants was derived from a power analysis expecting a medium effect for the main effect of 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. They were allocated to three different groups (motor experts, visual experts and novices) depending on their gymnastics expertise. This was done to control for the influence of both visual and bodily experiences on perception and judgment (Pizzera & Raab, 2012). The group of motor experts ($n = 17$) included persons (Women: $n = 10$; $M_{\text{age}} = 25.71$, $SD = 13.70$, range = 12–56), who were active in gymnastics as athletes for a minimum of three years and were still active at the date of the study. The group of visual experts ($n = 14$) included persons (Women: $n = 9$; $M_{\text{age}} = 31.86$, $SD = 14.26$, range

= 19–75), who were active in gymnastics as judges or trainers for a minimum of three years and were still active at the date of the study. The group of novices ($n = 18$) included persons (Women: $n = 9$; $M_{\text{age}} = 30.89$, $SD = 7.43$, range = 12–44), who were not active and have never been active before in gymnastics, neither as athletes nor as judges or trainers. The age, $F(2, 46) = 1.17$, $p = .321$, and gender, $H(1) = 0.017$, $p = .9$, did not differ significantly between the groups. The participants voluntarily took part in this study 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, presented in original and in 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 competition. This was done to have different performances and different surface information in the video stimuli. The camera was placed approximately 15 meters away from the place where the routines were performed, with its optical axis being orthogonal to the motion direction of the gymnast, simulating the judge's perspective. The stick figure videos were created by using the Simi Motion software. 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 (right and left sides of the body in each case), and head. The videos started with the gymnasts' last step of the take-off phase of the round off and ended shortly before the final landing position of the back layout somersault to avoid the influence of the final routine position on the judgment. Before the videos, a short video sequence showing the numbers 4, 3, 2, 1, and the symbol + in a sequence at an interval of 1 s was inserted at the left side of the screen where the floor routine starts. This was done to focus the participants' attention on the starting place and time of the first skill (see Figure 1).

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): (a) technically poor execution or fall, (b) technical errors with large postural errors, (c) technical errors with minor postural errors, (d) good execution with minor postural errors, (e) good execution, and (f) very good execution. The scale was presented on a tablet placed by each participant in a comfortable position in front of the screen. On the scale, the six points were marked, but intermediate stages ($n = 10$) could also be chosen. Each video was shown only once.

To calculate performance accuracy, the absolute difference of the judgment score to the true scores was calculated. To obtain the true

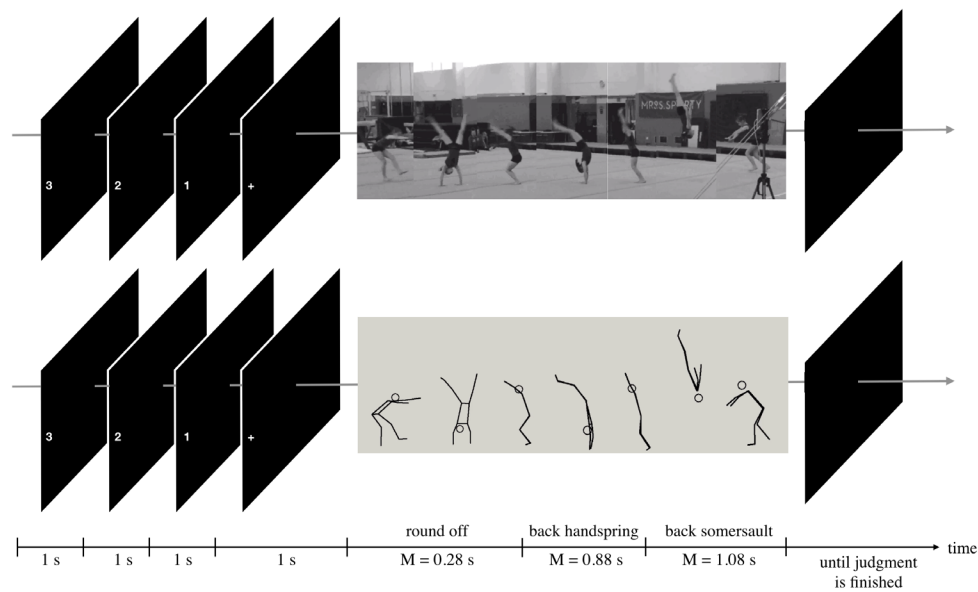


FIGURE 1.

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

scores, the same videos were rated on the same scale by three gymnastics experts (visual and motoric experience in research and praxis for over ten years). They were allowed to view the videos in their preferred pace, as often as they wanted, and to stop them whenever they wanted. 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 away 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 a participant fixated on a certain location on the screen, starting at a minimum of 50 ms (Galley et al., 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 durations over the total number of fixations. After removing trials where 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 whole 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 the gaze parameters which were shown, on average, in the first 13.09 cs ($SD = 1.20$ cm) of the screen

and the first 0.28 s ($SD = 0.01$ s) of the action duration. For the back handspring, this means the gaze parameters which were shown, 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 action duration. For the somersault, this means the gaze parameters which were shown, 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 last approximately 30 minutes and each participant was tested individually. At the beginning, the participant was informed about the procedure of 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 normal, which is typical for gaze data (Galley et al., 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 performed for the dependent variables of average fixation duration, summarized fixation 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 performed 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 as well as the judgment measurements (judgment score and accuracy) were assessed and analyzed because both aspects have their limitations as well as advantages, and thus complement each other. Concerning the judgment measurements, both the score as well as the accuracy was measured and analyzed because those are both 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 to facilitate insights on 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

Judgment

Table 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 of accuracy were assumed as normally distributed. Levene's test indicated equal variances for the

judgment score, $F = 1.23$, $p = .303$, and unequal variances for the judgment accuracy, $F = 2.58$, $p = .031$.

JUDGMENT SCORE

For judgment score, the ANOVA revealed a significant main effect of expertise, $F(2, 46) = 7.35$, $p = .002$, $\eta_p^2 = .242$, and format, $F(1, 46) = 13.27$, $p < .001$, $\eta_p^2 = .224$, but no significant interaction effect, $F(2, 46) = 0.71$, $p = .499$, $\eta_p^2 = .03$. 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 the Holm correction showed that both motor experts and novices rated the skills more highly than did visual experts ($p < .001$, see Figure 2).

JUDGMENT ACCURACY

For judgment accuracy, the ANOVA revealed a significant main effect of expertise, $F(2, 46) = 5.62$, $p = .007$, $\eta_p^2 = .196$, but no significant main effect of format, $F(1, 46) = 0.01$, $p = .941$, $\eta_p^2 < .001$, and no significant interaction effect, $F(2, 46) = 2.30$, $p = .112$, $\eta_p^2 = .048$. A post hoc t -test with the 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 ($p = .024$, see Figure 2).

Gaze Parameters

WHOLE FLOOR ROUTINE

Table 1 shows the descriptive statistics of the gaze parameters. To investigate the gaze patterns for the floor routines as a whole, three 3 (Group) \times 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 normally distributed,

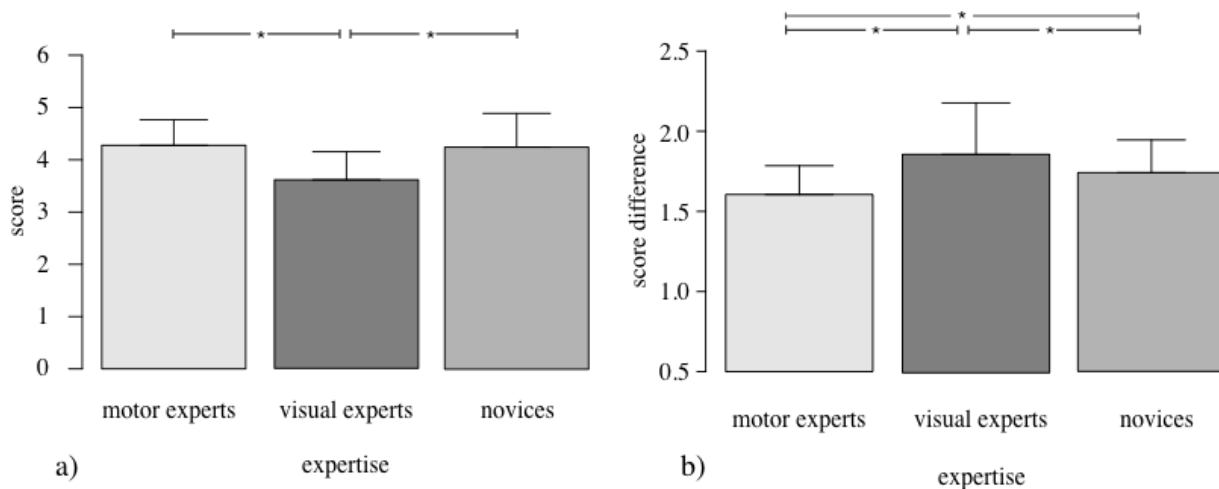
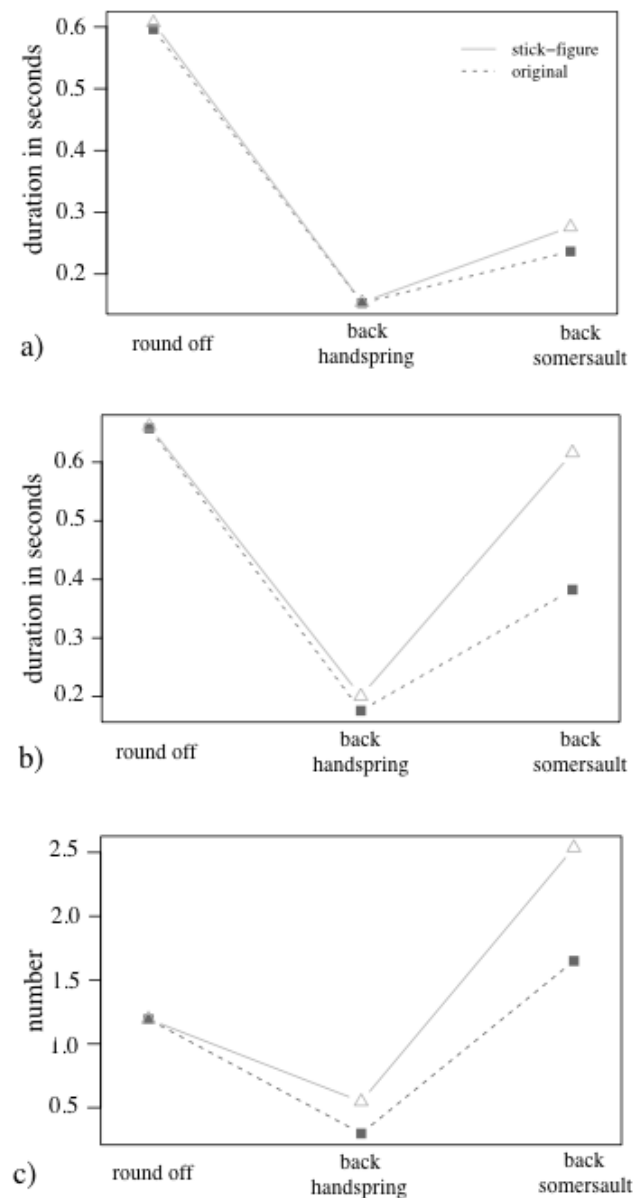


FIGURE 2.

Means and SDs of judgment scores (a) and the accuracy with regard to expertise (b). * = significant differences.

**FIGURE 3.**

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

which is typical for gaze data (Galley et al., 2015). The residuals of the summarized fixation can be assumed as 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 average fixation duration, the ANOVA revealed a significant main effect of format, $F(1, 46) = 17.00$, $p < .001$, $\eta_p^2 = .270$. All other effects were not significant ($p > .05$). Skills which were presented in the

stick-figure format led to a higher average fixation duration ($M = 0.484$, $SD = 0.154$) than skills that were presented in the original format ($M = 0.432$; $SD = 0.105$).

SUMMARIZED FIXATION DURATION

No significant effects were revealed by the ANOVA for summarized fixation duration ($p > .05$).

NUMBER OF FIXATIONS

For the number of fixations, the ANOVA revealed a significant main effect of format, $F(1, 46) = 17.51$, $p < .001$, $\eta_p^2 = .276$, and no other signifi-

TABLE 1.Means and *SD*s of Gaze Parameters as a Function of Format and Expertise for the Temporal-Spatial Routine Windows

| Gaze parameters video format | Average fixation duration | | | | Summarized fixation duration | | | | Number of fixations | | | |
|---------------------------------|---------------------------|-----------|--------------|-----------|------------------------------|-----------|--------------|-----------|---------------------|-----------|--------------|-----------|
| | Original | | Stick-figure | | Original | | Stick-figure | | Original | | Stick-figure | |
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> |
| 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 | 0.497 | 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 |

cant effects ($p > .05$). Skills which were presented in the original format led to a higher number of fixations ($M = 6.296$, $SD = 0.986$) than skills that were presented in the stick figure format ($M = 5.850$; $SD = 1.122$).

Temporal-Spatial Skill Windows

To investigate the gaze pattern for the floor routines, separated into the three different skills, three 3 (Group) \times 2 (Format) \times 3 (Skill) three-way repeated-measures ANOVAs were performed 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, $F(2, 92) = 32.07$, $p < .001$, $\eta_p^2 = .411$. A post hoc t -test with the Holm correction showed a significant difference concerning the summarized fixation duration for all of the three skills ($p < .05$). Thus, the back handspring leads to the lowest average fixation duration ($M = 0.154$, $SD = 0.335$), followed by the back layout somersault ($M = 0.257$; $SD = 0.154$), and the round off ($M = 0.601$; $SD = 0.422$). All other effects were not significant ($p > .05$, see Figure 3).

SUMMARIZED FIXATION DURATION

For the summarized fixation duration, the ANOVA revealed a significant main effect of skill, $F(2, 92) = 29.00$, $p < .001$, $\eta_p^2 = .387$, and format, $F(1, 46) = 11.67$, $p = .001$, $\eta_p^2 = .202$, as well as a significant interaction effect between skill and format, $F(2, 92) = 6.09$, $p = .003$, $\eta_p^2 = .117$ (see Figure 1). All other effects were not significant ($p > .05$). A post hoc t -test with the Holm correction showed a significant difference concerning the summarized fixation duration for all of the three skills ($p < .05$). Thus, the back handspring leads to the lowest summarized fixation duration

($M = 0.188$, $SD = 0.387$), 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 led to a higher summarized fixation duration ($M = 0.493$, $SD = 0.411$) 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 (see Figure 3).

NUMBER OF FIXATIONS

For the number of fixations, the ANOVA revealed both a significant main effect of skill, $F(2, 92) = 179.19$, $p < .001$, $\eta_p^2 = .796$, and format, $F(1, 46) = 53.69$, $p < .001$, $\eta_p^2 = .539$, as well as a significant interaction effect between skill and format, $F(2, 92) = 30.263$, $p < .001$, $\eta_p^2 = .397$ (see Figure 1). All other effects were not significant ($p > .05$). A post hoc t -test with the Holm correction showed a significant difference concerning the number of fixations for all of the three skills ($p < .05$). Thus, the back handspring led to the lowest number of fixations ($M = 0.422$, $SD = 0.724$), 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 led to a higher number of fixations ($M = 1.426$, $SD = 1.028$) than did skills 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 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 ($Z = 3.43, p = .904$). 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: $M = 0.382, SD = 0.101$; see Figure 3).

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 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) 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.

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 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 a whole 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. Additionally, in further 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 (Brybaert, 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.

ACKNOWLEDGEMENTS

The author would like to thank the gymnasts and the coaches of the gymnastics club "TuS 1861 Chemnitz-Altendorf e.V." for the opportunity to film the gymnastics movements, the coaches and athletes of the gymnastics club "TV Markkleeberg von 1871 e.V." for the rating of the videos, and all the other participants and persons who assisted the study. We acknowledge support from the German Research Foundation (DFG) and Leipzig University within the program of Open Access Publishing.

REFERENCES

Abernethy, B., Gill, D. P., Parks, S. L., & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception*, *30*, 233–252. doi: 10.1068/p2872

Abernethy, B., & Zawi, K. (2007). Pickup of essential kinematics underpins expert perception of movement patterns. *Journal of*

Motor Behavior, *39*, 353–367. doi: 10.3200/JMBR.39.5.353-368

Abernethy, B., Zawi, K., & Jackson, R. C. (2008). Expertise and attunement to kinematic constraints. *Perception*, *37*, 931–948. doi: 10.1068/p5340

Ansoerge, 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. American Alliance for Health, Physical Education and Recreation*, *49*, 399–405. doi: 10.1080/10671315.1978.10615552

Arkaev, L., & Suchilin, N. (2004). *How to create champions*. Meyer & Meyer.

Bard, C., Fleury, M., Carriere, L., & Halle, M. (1980). Analysis of gymnastics judges' visual search. *Research Quarterly for Exercise and Sport*, *51*, 267–273. doi: 10.1080/02701367.1980.10605195

Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics*, *23*, 145–152. doi: 10.3758/BF03208295

Brybaert, 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*, 16. doi: 10.5334/joc.72

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

Creswell, J. & Tashakkori, A. (2007). Differing perspectives on mixed methods research. *Journal of Mixed Methods Research*, *1*, 303–308. doi: 10.1177/1558689807306132

Cicchetti, D. V. (1994). Guidelines, criteria and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284–290. doi: 10.1037/1040-3590.6.4.284

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. doi: 10.3758/BF03337021

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*, 357–372. doi: 10.1037/0096-1523.4.3.357

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*, 934–941. doi: 10.1007/s12144-017-9572-1

Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, *22*, 15–22. doi: 10.1068/p220015

Dosseville, F., Laborde, S., & Garnczerzyk, C. (2013). Current research in sports officiating and decision-making. In C. Mohiyeddini (Ed.), *Contemporary topics and trends in the psychology of sports* (pp. 13–38). Nova Publishers.

Deutscher Turner-Bund. (2001). *Handbuch des Deutschen Turner-Bundes, Teil 1, Aufgabenbuch, Broschüre 1 - Gerätturnen weiblich* [Handbook of the German Gymnastics Federation, Part 1, com-

- pulsory exercises, brochure 1, female artistic gymnastics]. 3rd ed. Deutscher Turner-Bund Service GmbH. 2001.
- Farana, R., Uchytíl, 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*, 33–40. doi: 10.5507/ag.2015.005
- 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*, 49–57. doi: 10.5507/ag.2012.012
- Fédération Internationale de Gymnastique [FIG] (2017). 2017 - 2020 Code of Points. Women's artistic gymnastics. Retrieved from http://www.fig-gymnastics.com/publicdir/rules/files/en_WAG%20CoP%202017-2020.pdf
- 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). 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*, 523–552. doi: 10.1007/s10648-011-9174-7
- Hars, M., & Calmels, C. (2007). Observation of elite gymnastics performance: Processes and perceived functions of observation. *Psychology of Sport and Exercise*, *8*, 337–354. doi: 10.1016/j.psychsport.2006.06.004
- Hauw, D., Renault, G., & Durand, M. (2008). How do aerial freestyle 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*, 481–486. doi: 10.1016/j.jsams.2007.06.005
- 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*, 63–72.
- Huys, R., Smeeton, N. J., Hodges, N. J., Beek, P., & Williams, A. M. (2008). On the dynamical information underlying visual anticipation skill in tennis. *Perception and Psychophysics*, *18*, 1217–1234. doi: 10.3758/PP.70.7.1217
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211. doi: 10.3758/BF03212378
- 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. doi: 10.1123/jtpe.23.2.123
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*, 155–163. doi: 10.1016/j.jcm.2016.02.012
- MacMahon, C., & Plessner, H. (2008). The sport officials in research and practice. In D. Farrow, J. Baker & C. MacMahon (Eds.), *Developing sport expertise. Researchers and coaches put theory into practice* (pp. 172–192). Routledge.
- 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). 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*, 457–478. doi: 10.1123/jsep.29.4.457
- Mercier, H., & Heiniger, S. (2018). Judging the judges: Evaluating the performance of international gymnastics judges. *ArXiv*, 1807.10021.
- 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. doi: 10.2466/pms.2002.95.3.901
- Pizzera, A. (2012). Gymnastic judges benefit from their own motor experience as gymnasts. *Research Quarterly for Exercise and Sport*, *83*, 603–607. doi: 10.1080/02701367.2012.10599887
- 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*, 112–119. doi: 10.1080/02701367.2017.1412392
- Pizzera, A., & Raab, M. (2012). Perceptual judgments of sports officials are influenced by their motor and visual experience. *Journal of Applied Sport Psychology*, *24*, 59–72. doi: 10.1080/10413200.2011.608412
- Plessner, H., & Haar, T. (2006). Sports performance judgments from a social cognition perspective. *Psychology of Sport and Exercise*, *7*, 555–575. doi: 10.1016/j.psychsport.2006.03.007
- Plessner, H., & Schallies, E. (2005). Judging the cross on rings: A matter of achieving shape constancy. *Applied Cognitive Psychology*, *19*, 1145–1156. doi: 10.1002/acp.1136
- 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*, 630–639. doi: 10.1016/j.psychsport.2012.03.008
- Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann, & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167–201). Springer
- Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*, 129–154. doi: 10.1080/713752551
- Ste-Marie, D. M. (1999). Expert–novice differences in gymnastic judging: An information-processing perspective. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *13*, 269–281. doi: 10.1002/(SICI)1099-0720(199906)13:3<269::AID-ACP567>3.0.CO;2-Y
- Ste-Marie, D. M. (2000). Expertise in women's gymnastic judging:

- An observational approach. *Perceptual and Motor Skills*, 90, 543–546. doi: 10.2466/pms.2000.90.2.543
- 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, 420–426. doi: 10.1080/02701367.2001.10608979
- Stevenage, S. V., Nixon, M. S., & Vince, K. (1999). Visual analysis of gait as a cue to identity. *Applied Cognitive Psychology*, 13, 513–526. doi: 10.1002/(SICI)1099-0720(199912)13:6<513::AID-ACP616>3.0.CO;2-8
- 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. doi: 10.1080/07303084.1987.10603894
- 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. doi: 10.1123/ijab.8.2.87
- Takei, Y. (1998). Three-dimensional analysis of hand-spring with full turn vault: Deterministic model, coaches' beliefs, and judges' scores. *Journal of Applied Biomechanics*, 14, 190–210. doi: 10.1123/jab.14.2.190
- Takei, Y. (2007). The roche vault performed by elite gymnasts: Somersaulting technique, deterministic model, and judges' scores. *Journal of Applied Biomechanics*, 23, 1–11. doi: 10.1123/jab.23.1.1
- Takei, Y., Blucker, E., 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, 849–863. doi: 10.1080/026404100750017788
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, 2, 371–387. doi: 10.1167/2.5.2
- Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics*, 67, 667–675. doi: 10.3758/BF03193523
- Ward, P., Williams, A. M., & Bennett, S. J. (2002). Visual search and biological motion perception in tennis. *Research Quarterly for Exercise and Sport*, 73, 107–112. doi: 10.1080/02701367.2002.10608997
- 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. doi: 10.1016/j.humov.2005.06.002
- Williams, A. M., Davids, K., & Williams, J. G. (1999). *Visual perception and action in sport*. E & FN Spon.
- Williams, A. M., Huys, R., Cañal-Bruland, R., & Hagemann, N. (2009). The dynamical information underpinning deception effects. *Human Movement Science*, 28, 362–370.

RECEIVED 12.07.2019 | ACCEPTED 06.12.2019