HOW DOES AUDITORY INFORMATION INFLUENCE OBSERVERS’ PERCEPTION DURING THE EVALUATION OF COMPLEX SKILLS?

Frederike Veit
Faculty of Sport Science, Leipzig University, Germany

Abstract
Perceiving and gathering information from the environment are essential abilities of humans, especially in sports. An individual’s perception of aspects such as the surrounding conditions or the movements of other athletes can be decisive for a successful performance. The question that arises is whether an individual’s level of expertise affects his or her ability to use task-specific information. Furthermore, it should be determined whether the type of information that is gathered differs across people with different sport-specific experiences. The aim of the study was to investigate the role of auditory information in the observation and evaluation of complex skills in gymnastics. Participants with different amounts of experience were asked to estimate the duration of flight phases in straight-back somersaults on the floor under manipulated auditory conditions. The results of the current study show that participants with no specific experience in gymnastics performed worse than participants with visual or motor experiences. Additionally, the current gymnasts outperformed the other participants. One could speculate that current gymnasts benefit from their motor experiences, which lead to improved perceptual sensitivity and a better ability to identify differences between two cues. In conclusion, it could be enriching to take the auditory information into account in motor learning tasks.

Keywords: common-coding, paired comparison, perceptual sensitivity, acrobatic series on floor, expert-novice paradigm.

INTRODUCTION
Perceiving and gathering information from the environment are essential abilities of humans, especially in sports. An individual’s perception of aspects such as the surrounding conditions or the movements of other athletes can be decisive for a successful performance. More specifically, the ability to use domain- and task-specific information may be a crucial advantage for athletes in generating an effective motor action (Mann, Williams, Ward, & Janelle, 2007). Not only the performance of athletes but also the performance of judges and coaches depend on their perceptual abilities. Recent studies have noted that the performance of judges and coaches relate to aspects such as their gaze behavior (Pizzera, Möller, & Plessner, 2018), their own motor experiences (Pizzera, 2012), and years of practical experience (Pizzera & Raab, 2012). The question that arises is whether an individual’s level of expertise affects his or her ability to use task-specific information. Furthermore, it should be determined whether the type of information that is gathered differs across people with different sport-specific...
experiences. Therefore, the aim of the study was to investigate the role of visual and auditory information in the observation and evaluation of complex skills in gymnastics. More specifically, the aim was to extend the knowledge on the extent to which people with different degrees of expertise rely on various sources of information.

According to the common coding theory, perceived information from the visual and auditory system activates the same representation that is utilized when executing a perceived action (Prinz, 1997). Hommel (2019) supplemented this general assumption with a conjecture about the degree of influence of the task-relevant dimension (e.g., technical characteristics). He describes that “codes of feature on a dimension that is (assumed to be) relevant for the presently relevant task will have a stronger impact on representing an event then codes of features related to currently irrelevant dimensions […]” (Hommel, 2019, p. 2141). Observing and evaluating skills in sports are of great importance.

In technical sports such as gymnastics, visual information is also the subject of the evaluation for judges. Therefore, the final ranking of the athletes depends on the perceptual analysis of judges (Ste-Marie, 1999). Relatedly, some studies have investigated the role of visual information in and which factors determine the judging performance and the quality of judgments (Heinen, Vinken, & Velentzas, 2012; Pizzera, Möller, & Plessner, 2018).

Heinen, Vinken, and Velentzas (2012) examined, for instance, the role of visual and motor experiences in judging performances. They asked gymnastic judges (only visual experiences) and laypeople (only motor experiences) to rate handsprings performed on vault. The results indicate that both visual and motor experiences can influence judging performance. Regarding visual experiences, Pizzera, Möller, and Plessner (2018) investigated, whether gaze behavior affects judging performance. Furthermore, they contextualized the gaze behavior with additional motor experiences. In total, they asked 35 judges to judge 21 gymnasts on vault via video performances. The participants differed in their expertise. The authors compared the performance and gaze behaviors of the judges with a high-level of expertise and a lower level of expertise (as measured by license levels) and the performance and gaze behaviors of judges, who can perform the skill on the vault and those who do not have the ability. Contrary to the results in previous studies, the results showed that specific motor experiences do not influence judging performance. Furthermore Pizzera, Möller and Plessner (2018) identified the certification level as an influencing factor.

Concerning the influence of different factors on perceptual judgments and judgment performance, in addition to the abovementioned factors (i.e., motor experiences, license level, years of judging experiences), Plessner and Schallies (2005) proposed that different factors related to the viewing position influence the evaluation of gymnastics skills. They asked 80 participants (40 male judges, and 40 students (26 female and 14 male) as layout people) to judge the cross on rings from three different viewpoints. They examined whether the viewing position of judges and laypeople, as the perceptual factor, influences the judging performance. The results showed that the judges performed much better than did the laypeople. Additionally, the judges were significantly more influenced by the viewing position than were the laypeople.

Additionally, in gymnastics, the auditory information from the natural movement seems to be an essential source of information (Veit & Heinen, 2019). Here different groups (i.e., current gymnasts, former gymnasts, judges/coaches, and laypeople) were asked to observe and evaluate a complex gymnastic skill in visual, audio-visual, and auditory conditions. The main focus of the study was on the judgment of the duration
of the flight phase of a straight-back somersault. The results showed that current gymnasts have a significantly higher score in the audio-visual and auditory condition than in the visual condition. Experienced athletes can use movement sounds to evaluate and to improve the quality of their performance (Agostini, Righi, Galmonte, & Brino, 2004; Stauffer, Haldemann, Troche, & Rammsayer, 2012). Judges/coaches, in contrast, outperformed the current and former gymnasts and laypeople in the visual condition.

There exists the assumption that different sensory information affects perception. Kennel, Streese, Pizzera et al. (2015) investigated the influence of natural movement sounds on motor control. More specifically, they conducted a study with 20 participants, who performed a hurdling task in three auditory conditions (normal movement sounds, manipulated movement sounds, and delayed movement sounds). The movement sounds were presented as feedback during the task through headphones. The results suggested that movement sounds, as a result of contacting the floor during running, have an impact on motor control.

Heinen, Koschnick, Schmidt-Maaß, and Vinken (2014) examined the role of auditory and visual information in skill performance, particularly during a synchronization task on trampolines. The experimental task of two gymnasts was to perform synchronized leaps. The study aimed to investigate the role of auditory and visual information in the synchronization of trampolining. To synchronize their performances, gymnasts must perceive and identify relevant information. Twenty female gymnasts had to perform straight leaps in three conditions (1. full visual and auditory information available, 2. limited auditory information available, 3. limited visual information available). Results showed that neither the visual information alone nor auditory information alone can explain the synchronization process. Nonetheless, it seems that visual information plays a dominant role, which is supported by auditory information.

The studies described above illustrate that perception is driven by different sensory information. The state-of-the-art shows that the main focus is on visual aspects. However, an increasing number of recent studies have been conducted on other sources of information, such as auditory cues, which also provide important information for action perception and movement execution. In the context of motor learning, the role of auditory information seems to play an important role. Baudry, Leroy, Thouvaréq, and Chollet showed in 2006 that auditory information during the learning process could be helpful for gymnasts. One could speculate that focusing on the auditory information (e.g., rhythm) makes it easier to develop a representation of the movement. In the track and field, previous studies have highlighted the significance of auditory information and feedback for motor learning, action perception and agent identification (hammer throwers, Agostini, Righi, Galmonte, & Bruno, 2004; hurdling, Kennel, Streese, Pizzera et al., 2015).

It is common ground that different sensory information affects perception. In recent studies, primarily the visual information has been examined. Moreover, the research focus has moved to the auditory sources in different contexts and sports (complex movements) and various domains (e.g., anticipation, agent identification, motor learning). However, it remains unknown how the auditory information is used by people with different sport-specific experiences in gymnastics.

It was hypothesized that the quality of observation and evaluation in gymnastics depends on the expertise of the participants (Heinen, Vinken, & Velentzas, 2012). Overall, it was expected that gymnastics experts (judges/coaches as visual experts and current gymnasts as...
motor experts) outperform laypeople. It was expected that the manipulated auditory sounds do not influence judges/coaches (visual experts) and that they make decisions primarily based on the visual information rather than the auditory information, even when the auditory information indicates conflicting information caused by the manipulation of the audio file. In contrast, it was expected that the current gymnasts make decisions based on auditory information, even when the visual information indicates conflicting information (Veit & Heinen, 2019).

METHODS

In total, $N = 36$ participants took part in this study (age range 19-36; $M_{age} \pm SD = 28 \pm 4$ years). The recommended number of participants was derived from a power analysis (Cohen’s $f^2 = 0.25$, type-I-error probability $= 5\%$, type-II-error probability $= 20\%$). The participants were assigned to one of three groups based on their motor and visual experience and expertise in gymnastics: group 1 - neither motor nor visual expertise with the experimental task ($n_1 = 12$ laypeople), group 2 - extensive visual experience with the motor task ($n_2 = 12$ judges/coaches), and group 3 - extensive motor expertise with the motor task at the time of the study ($n_3 = 12$ current gymnasts).

The laypeople were defined as people having no or only minor experiences with watching or performing gymnastics. The judges and coaches had substantial experience (at least seven years) in observing, coaching and judging gymnastics skills in general and with the experimental task used in this study in particular. The current gymnasts (motor experts) were able to perform the acrobatic series (see below) on the floor at the present time and had at least ten years of gymnastics experience in general, particularly in performing a round-off, followed by a back handspring and a straight-back somersault on the floor.

None of the participants reported having any hearing disorders or corrected-to-normal vision. They were furthermore informed about the procedure and aim of the study and gave their written consent before the beginning of the study. The study was conducted in line with the ethical guidelines of the local ethics committee.

The experimental task of the study was to observe an acrobatic series on the floor. The series comprised three gymnastics elements, namely a round-off, followed by a back handspring, and a straight-back somersault (George, 2010). The series was performed on a standard gymnastics spring floor, and these elements constitute a common acrobatic series in floor exercises (Arkaev & Suchilin, 2004). While the round-off followed by a back handspring function as preparatory elements (i.e., the generation of angular momentum towards the take-off of the somersault), the straight-back somersault is usually characterized by maximization of flight duration and an optimization of angular momentum (Prassas, Kwon, & Sands, 2006). During the flight phase, gymnasts perform a rotation of 360 degrees in a straight body posture about the somersault axis before landing on the ground.

Prior to data acquisition, ten female expert gymnasts were asked to perform the acrobatic series six times. Their performances were videotaped with a Full-HD digital video camera recording data at 240 Hz. The camera was placed orthogonal to the execution direction at a distance of approximately 15 meters. From the recorded video sequences, six were selected for further stimulus generation. The selection was made based on there being little background noise and clear natural thumping sounds when the individuals made contact with the floor. The thumping sound was the result of the
gymnasts contacting with the floor when performing the acrobatic series.

The recorded video sequences were digitized frame by frame using the movement analysis system Simi Motion® version 9 (Simi Reality Motion Systems GmbH, Munich, Germany). Coordinates of 14 body landmarks corresponding to the joints of a ten-segment model of the human body were digitized for each frame of the video sequences (Enoka, 2008). For each digitized video sequence, a stick-figure video sequence was generated, consisting of a black stick-figure in front of a gray background and a straight horizontal line that represented the floor (see Figure 1 as an example). This was done to mask the personal characteristics of the gymnasts, making it easier for the participants to focus on the movement information instead of the surface characteristics, such as the leotard color and background objects. In addition, the stick-figures were scaled to the same size. The audio information from the original video file was added to the stick-figure video sequences.

Figure 1. Stick-figure sequence of the motor task (round-off, back handspring and straight-back somersault).

From each original stick-figure video sequence, two additional and modified versions were created using the software iMovie (version 10.1.14, © 2001-2019 Apple Inc.). The natural movement sound from the take-off of a straight-back somersault on the floor was manipulated (+23%, -23% loudness) to generate the different conditions. In one version, the loudness of the original audio information was increased by 23%, while in the second version, the loudness was decreased by 23%. Twenty-three percent was chosen because it was the percent of change that was not noticeable according to previous pilot-studies and expert evaluations. At the
end of the aforementioned process, every stick-figure sequence was available in three different versions (original, +23% loudness, -23% loudness).

For example, in the unchanged version, one of the longest somersaults had a flight duration of 0.77 seconds and a short-term loudness during the take-off of -29.6 LUFS (loudness unit), while one of the shortest somersaults had a flight duration of 0.62 seconds and a short-term loudness during the take-off of -34.7 LUFS.

The participants’ task was to observe stick-figure video sequences of the acrobatic series (described above). The stick-figure sequences were presented in pairs on a 27-inch computer screen with one sequence in the upper part of the computer screen and the other sequence in the lower part of the screen (see Effenberg, 2005, and Kennel et al., 2015 for a similar approach). Auditory information for the stick-figure sequences was played via a Sony MDR-XB950AP headphones (Sony Corporation, Tokyo, Japan) that the participant wore while completing the task. The experimental task was to correctly identify the straight-back somersault video with the longer flight phase. Thus, after watching the two stick-figure videos, the participants had to select one of the three response options: 1. The upper stick-figure sequence comprised a straight somersault with the longer flight phase than did the bottom one, 2.) the bottom stick-figure sequence comprised the straight somersault with the longer flight phase, or 3.) the upper and lower stick-figure sequence comprised a straight somersault with the same flight duration. The second video-clip was presented two seconds after the first clip, and the next pair of stick-figure sequences were presented after three seconds. In total, the test lasted approximately 25 minutes.

In the current study, five conditions were realized: 1) baseline (auditory information and the video file were unchanged), 2) the loudness of the audio file of the video clip with the longer flight duration was increased by 23% (longer +23%), 2a) the loudness of the audio file of the video-clip with the shorter flight duration was increased by 23% (shorter +23%), 3) the loudness of the audio file of the video clip with the longer flight duration decreased by 23% (longer -23%), and 3a) the loudness of the audio file of the video-clip with the shorter flight duration was decreased by 23% (shorter -23%). For the baseline pairs, the longer, the louder-assumption clearly applied. Thus, the video clip with the longer flight duration in the layout somersault always corresponded to the clip with the louder auditory information. However, in conditions 3 and 2a in particular, the aforementioned manipulation of auditory information resulted in conflicting information (i.e., a longer flight phase/lower loudness or shorter flight phase/higher loudness compared to another stick-figure video sequence).

To prevent the position in which the video-clip is presented from affecting the results, every item was presented in the upper part of the computer screen as well as in the lower part. Additionally, the order of the conditions and the order of the trials were randomized. In total, 75 trials were presented to each participant. The center of interest was the participants’ choice. For every participant, the ratio of correct decisions was calculated. The ratio of correct decisions per condition was calculated as the sum up of the correct answers from each trial divided by the number of trials per condition. For each condition, every participant was able to make a maximum of 15 correct decisions per condition and a minimum of 0. The participants’ answers were recorded by an online questionnaire that they completed on a tablet (see below).

In order to collect the data, an online questionnaire was created. It consisted of a general section comprising demographic questions and a specific section comprising questions related to the participants’
specific experiences in gymnastics. In addition to the participant’s age and sex, his or her current activity in gymnastics was addressed by the questions. The information about the current activity was addressed in four sections within the specific section of the questionnaire. The current gymnasts had to answer four more questions about their experiences in gymnastics (years that they have been actively practicing, weekly training hours/frequency/amount of training, participation, and level of competitions attended). Coaches and judges had to give information about the amount of coaching or judging they have performed in hours, the level of competitions attended, and whether the floor sequence is an integral part of their training or competitions. At the end of the questionnaire, every participant answered the question on whether they ever performed the sequence on the floor. The third part of the questionnaire was presented to the participants during the experimental task so that the participants’ answers for each experimental trial could be recorded.

The study was conducted in three phases and lasted approximately 35 minutes. In the first phase, the participants arrived at the laboratory and were informed about the procedure, the task, and the general purpose of the study. Afterwards, the participants used a tablet to complete the first two parts of the questionnaire, comprising general sociodemographic questions as well as specific questions about his/her experiences in gymnastics as a gymnast, as a coach, or as a judge. After the questionnaire, the participants were shown several exemplary stick-figure sequences so that they were familiarized with the task. After familiarization, the gymnasts had the opportunity to ask some questions.

In the second phase, data acquisition was conducted. After watching a trial, each participant had to choose one of three response options: 1.) the upper video clip showed the longer flight phase or 3.) the upper and lower video clips flight phases of the same duration. In total, he/she had to make 78 decisions. At the beginning of the data collection, three additional pairs were shown for the familiarization (see above), and these pairs were not included in the analysis. The subsequent 75 decisions were used for further data analysis. After 40 trials, the participants were given a short break to prevent a loss of concentration.

In the third phase, the effect of manipulation check was verified by asking the participant whether he/she noticed anything abnormal during the data collection. No data had to be excluded because there were no positive response to the manipulation check (meaning none of the participants noticed the sound was manipulated). Finally, he/she received a reward for participating in the study.

The significance level was set at 5% for all results. Before the main hypotheses were tested, whether the data met the assumptions for analysis of variance were determined. The results of the Kolmogorov-Smirnov test and the Mauchly test indicated that the data could be considered normally distributed and that the assumption of sphericity was not violated.

RESULTS

It was hypothesized that the quality of observation and evaluation in gymnastics depends on the expertise of the participants. Overall, the gymnastics experts (judges/coaches as visual experts and current gymnasts as motor experts) were expected to outperform the laypeople. Given the aforementioned hypothesis, it was expected that the manipulated auditory sounds would not influence the decisions of the judges/coaches. They should base their decisions primarily on visual information, even when auditory and visual information is conflicting (e.g., the auditory information is manipulated).
However, the current gymnasts should base their decisions first and foremost on auditory information, even when the visual and auditory information are conflicting. Additionally, it was hypothesized that increasing the loudness of the videos can increase the number of correct decisions.

![Figure 2](image-url)

*Figure 2. Means and standard errors of the participants’ relative number of correct decisions in the experimental conditions.*

To test the hypotheses, an analysis of variance (ANOVA) with repeated measures was performed. The decision ratio was included as the dependent variable. The ANOVA results show a significant main effect of condition, $F(4, 132) = 3.143$, $p = .017$, Cohen’s $f = 0.31$, as well as a significant main effect of group, $F(1, 3) = 9.375$, $p < .001$, Cohen’s $f = 0.75$. There was no significant interaction effect, $F(8, 132) = .874$, $p = .540$. The results suggest different decision ratios across the conditions and between the groups. The post hoc tests demonstrate that the participants made more correct decisions in the baseline condition than in the other conditions. Interestingly, a close examination of the decision ratios of the groups within the conditions shows that the current gymnasts did not make fewer correct decisions regardless of whether the loudness was increased or decreased. Between groups the ratio of correct decisions of the laypeople was significantly lower than that the current gymnasts. Additionally, there was no significant effect between the judges/coaches and current gymnasts ($p = .075$). A study with a larger sample size should be conducted to determine whether this effect is significant at the significance level of 5%.

**DISCUSSION**

The aim of the current study was to investigate the role of visual and especially auditory information in the observation and evaluation of complex skills in
gymnastics. The participants had to evaluate the duration of the flight phase of somersaults as the last element in an acrobatic series on the floor. The participants differed in their expertise in gymnastics (laypeople with no experience in gymnastics, coaches/judges as visual experts, current gymnasts as motor experts). The natural movement sound from the take-off of a straight-back somersault on the floor was manipulated (+23%, -23% loudness), to generate different testing conditions.

The results show that participants without any specific experiences in gymnastics performed worse than did the participants with visual or motor experiences, which is in line with the results of previous studies (Pizzera & Raab, 2012). In total, the ratio of correct decisions in the evaluation of the duration of straight-back somersaults was the highest in the current gymnasts (motor experts), but only the difference between the laypeople and current gymnasts was significant. Regarding whether there is a weighting of motor or visual expertise in the evaluation of complex skills, a conclusion cannot be made from these results. The ratios of correct decisions illustrate that specific experiences are essential, but neither the visual experts nor the motor experts had a significantly different number of correct decisions. If there was be a significant difference in the decision ratios between the current gymnasts and the judges/coaches, one would be able to identify a weighting of motor or visual expertise.

Following the common-coding idea of Prinz (1997), perception and action possess shared representations. One could speculate that current gymnasts benefit from their motor experiences, which lead to improved perceptual sensitivity and a better ability to identify differences between two cues. When taking into account the above calculated statistical trend, the current gymnasts did outperform the visual experts. Regarding the conditions, it is noteworthy that the quality of observation and evaluation of the duration of straight-back somersault were relatively consistent.

To determine whether there is a weighting of visual or auditory information, the auditory information was manipulated. There were no significant differences between the groups within conditions 2 and 2a (increased loudness). One could speculate that the auditory information strengthens the differences between two flight phases, which makes it easier for participants to identify discrepancies independent of their experience level, leading to more correct decisions. More specifically, auditory information supports visual information (Eysenck & Keane, 2010). It is noteworthy that current gymnasts significantly outperformed the laypeople (and significant tendency to the judges/coaches) within condition 3 and condition 3a (decreased loudness). This result might, in part, be explained by the auditory information ceasing to support the visual information, making it more challenging to identify the differences between two cues; thus, the participants had to rely on their experiences. According to the common-coding theory (Prinz, 1997), the common codes of motor experiences and perceptual experiences may be useful for the observation and evaluation of the duration of straight-back somersaults. It seems that the motor experiences of the current gymnasts are supportive for the observation and evaluation of the duration of straight back somersaults. Additionally, conditions 3 and 2a included trials with conflicting information, so perceptual sensitivity was challenged when the participants evaluated the flight durations. However, there may be a task-specific effect because the current gymnasts were able to perform the acrobatic series at the time of the investigation.

It is acknowledged that this study has several limitations, and two aspects should be highlighted. First, gymnasts, judges and
coaches usually do not have to judge or evaluate stick figures, which were used as experimental stimuli in this study. However, stick figures were used to control the surface characteristics, which is nearly impossible when actual video sequences are used. Nevertheless, one could speculate that the unusual stimuli may have influenced the decision ratios. Nevertheless, none of the participants reported being disturbed by the video format used in the current study.

Second, judges (especially in female gymnastics) usually do not hear the natural movement sound as precisely as it was presented to them in the study. However, this was done to evaluate the role of auditory information when they observed gymnastics skills. Nevertheless, additional studies on the different components of the auditory information that correspond to the gymnastics apparatuses (i.e., amplitude, frequency) should be conducted.

CONCLUSION

According to the current results, one could conclude that current gymnasts weigh visual and auditory information differently than do judges/coaches. Therefore, it may be enriching to complement visual/verbal feedback with auditory feedback (Kennel, Streese, & Pizzera et al., 2015). Therefore, sensitization for gymnasts (the use of auditory information) and coaches (for different learning types) may be beneficial.

Additionally, judges of women’s gymnastics routines on the floor do not typically hear natural movement sounds (depending on the competition area, level, etc.). Furthermore, it would be interesting to determine whether a “perfect fit” of music and acrobatic or gymnastic series on the floor influences the perception and, consequently, the evaluation of complex skills in gymnastics. A potential task may be to ask judges to evaluate different series on the floor (acrobatic series and gymnastic jumps and leaps) that are additionally manipulated. There exists the possibility that increased or coordinated music during the takeoff or reactive leap of a somersault by a kick drum, for example, influences judges’ perception of flight phases or movement amplitudes.

ACKNOWLEDGEMENTS

The author thanks the judges, coaches, gymnasts and laypeople for participating in the study.

REFERENCES


Simi Motion ® (Version 9) [Computer software]. Munich, Germany: Simi Reality Motion Systems GmbH.


**Corresponding author:**

Frederike Veit  
Leipzig University  
Faculty of Sport Science  
Jahnallee 59, 04109 Leipzig, Germany  
e-mail: frederike.veit@uni-leipzig.de  
phone number: +49 (0) 341 97-31709