Journal of Experimental Psychology: Applied
2007, Vol. 13, No. 3, 158–170

Copyright 2007 by the American Psychological Association
1076-898X/07/$12.00 DOI: 10.1037/1076-898X.13.3.158

Expertise-Based Differences in Search and Option-Generation Strategies

Markus Raab
University of Flensburg

Joseph G. Johnson
Miami University

The current work builds on option-generation research using experts of various skill levels in a realistic task. We extend previous findings that relate an athlete's performance strategy to generated options and subsequent choices in handball. In a 2-year longitudinal study, we present eye-tracking data to independently verify decision strategies previously inferred from patterns of generated options. A verbal protocol identified the option-generation process for each individual prior to an allocation decision. Although athletes of varying expertise generated the same number of options on average, these options differed in quality between expert, near-expert, and nonexpert athletes for both their initial and final choices. These and other key results are formalized to elaborate a model of option generation, deliberation, and selection.

Keywords: eye-tracking, sports, decision making, longitudinal study, learning

At first glance, making decisions under conditions of high pressure, limited time, and restricted resources seems to be a complex task. For the experienced athlete, however, no intricate calculation seems to be involved; rather, the best option often just comes to mind. For instance, Therese Brisson (2003), a 2002 Olympic gold medalist in hockey, put it this way: "There is no time in hockey to evaluate all options and pick the best one. You have to choose the first, best one" (p. 217). Such seemingly intuitive behaviors (decisions) by experts are not outside the realm of formal modeling. That is, even if the observable behavior seems more intuitive than algorithmic—even if the expert's own metacognitive perception is of a "natural" or "reflexive" behavior—it is still possible to understand the mental processes involved. The overall goal of the present research was to develop a model that captures the mental processes that occur from the presentation of a decision situation to the selection of a course of action. Specifically, we wanted to describe the link between the use of different information search strategies, the subsequent option-generation process, and the resulting choice characteristics in a realistic sports task.

In the field of sport sciences, research on expertise in tactical decision making is quite limited (McPherson & Kernodle, 2003; see also articles in the special issue of *Psychology of Sport & Exercise* focusing on decision making edited by Bar-Eli & Raab, 2006). Furthermore, the components of decision-making processes in many real-world settings are not well understood. In particular, our knowledge of how experts of different levels assess a situation, generate options, and choose among the options generated is still incomplete. One major concern is that there are no longitudinal investigations concerning both decision-making and eye-tracking research in human behavior (McMorris, 1999). The purpose of the current research was to attempt to answer some of these open questions by recruiting handball athletes with varying levels of expertise, recording both outcome and process measures across the entire experimental trial and including multiple sessions for each participant. We wanted to capture and quantify expert performance and understand the mechanisms that differentiate expert, near-expert, and nonexpert performers in these complex choice situations. It is one thing to say that experts and nonexperts differ, but it is far more productive to understand why and how they differ. In particular, we focus on three distinct processes that serve to guide the ultimate selection of a course of action—information search, option generation, and deliberation producing choice.

To better understand these processes, we employed a combination of process-tracing methods that included both performers' gaze behavior (via eye tracking) and their option-generation procedure (via verbal protocol), in addition to recording their resulting choices and reaction times as outcome measures (for a similar procedure in the domain of electrical circuit troubleshooting, see van Gog, Paas, & van Merrienboer, 2005). The simultaneous collection of these variables greatly expands the scope of previous investigations in dynamic sport domains (Martell & Vickers, 2004). Moreover, no longitudinal studies in which researchers have captured these issues together have been reported so far. Before presenting the study we provide a short overview of existing research involving these cognitive processes in tactical team sport situations, paying particular attention where the extant literature addresses expertise effects.

Information Search (Eye Gaze)

Eye gaze has been used as a proxy for visual attention in information search (Rayner, 1998) in a great deal of previous research in applied domains. We limit our overview of gaze

Markus Raab, Institute for Movement Science and Sport, University of Flensburg; Joseph G. Johnson, Department of Psychology, Miami University.

The work described in this paper was supported by grants from the Alexander von Humboldt Foundation (TransCoop Raab/Johnson, 2006–2008) and the BISp (German Federal Institute for Sport Science, Raab VF 070866/04). We thank Hilke Zastrow for help in data collection; Maik Jepsen, Klaus Gärtner, and Arne Warnke for data analyses; Anita Todd for editing the manuscript; and Uwe Czienskowski for statistical discussions. We thank the reviewers Gary Klein, Henning Plessner and an anonymous reviewer for their constructive comments.

Correspondence concerning this article should be addressed to Markus Raab, University of Flensburg, Institute for Movement Science and Sport, Auf dem Campus 1, 24943, Flensburg, Germany. E-mail: raab@uni-flensburg.de

characteristics to their role in ball games such as the one employed in the current study (for general overviews in sports, see Williams, Davids, Burwitz, & Williams, 1992; Williams, Janelle, & Davids, 2004). Visual attention can provide important insights to the information used in decision situations, such as the importance of various attributes. For instance, in soccer penalty shooting it is first important for the goalkeeper to look at the approach of the player toward the ball and later to his shooting foot (Savelsbergh, Williams, van der Kamp, & Ward, 2002).

A common finding is that elite players rely on fewer gaze fixations, but of longer duration and to more important areas of the visual field, compared to novices (for an overview, see Williams, Davids, & Williams, 1999). However, there are also findings that show more fixations by experts than by novices (Williams, Ward, Knowles, & Smeeton, 2002). It is difficult to draw general conclusions because these studies have been conducted across many qualitatively different domains (i.e., different sports). There are also several differences in experimental design, including stimulus presentation times that vary from very few to several seconds, and decision times have varied from less than a second (Helsen & Pauwels, 1993) to 9.5 s (Williams et al., 1994). The absence of a standard definition of experience is an additional difficulty in determining aggregate effects of expertise. Due to the small number of participants in each group—ranging from 5 (Bard & Fleury, 1976) to 15 (Helsen & Pauwels, 1993) at most—effect sizes are an appropriate comparison standard. Despite the above domain and design differences, large effects are almost exclusively reported if experts and novices are compared, and to a lesser extent if experienced and less-experienced players are compared (e.g., see Williams & Davids, 1998).

The majority of researchers have reported single measures that summarize gaze characteristics across the entire course of a single trial. One notable exception is the work of Martell and Vickers (2004), who examined expert choices under high time pressure in hockey. The results of Martell and Vickers show that for more experienced players in a defense situation, gaze duration increased progressively as a play developed, with a stable gaze (or "quiet eye") preceding action. Longer periods of stable fixations before movement execution were also interpreted as "programming time" or stable information-movement coupling (for further explanations, see Williams et al., 2004). These results show differences in the gaze characteristics between elite and near-elite hockey players, but they do not relate such differences in gaze directly to subsequent cognitive processing, such as option generation.

Option Generation

In many applied situations, one is not presented explicitly with a set of options from which to choose. Although this may be the case in experimental decision-making tasks, in many real settings one must generate possible solutions to a problem, in which there may be more than one "correct" or satisfactory answer. Surprisingly, little work has been done on this topic; the work that has been done has focused primarily on the number and type of generated options, with less emphasis on the information search characteristics or subsequent choice from among the generated set.

In earlier work we developed a heuristic for selecting an option from a number of those self-generated. Like most simple heuristics, this so-called Take-the-First heuristic (TTF heuristic; Johnson & Raab, 2003) capitalizes on the extensive experience of the decision maker in the relevant environment. In particular, the TTF heuristic assumes that options are sequentially generated based on learned association strength between the candidate options and the current situation. Options that are repeatedly used in previous similar situations develop stronger links, and the TTF heuristic simply "bets" that these earlier-generated options are better. Furthermore, options seem to be stored in memory in a way that retains their functional attributes (e.g., whether an option is a pass or a shot on goal) and their spatial orientation (e.g., whether an option is on the left or right of the playing field). As a result, each option is associated with other options by virtue of both its functional and spatial properties. These connections can also determine options generated through "spreading activation," rather than through direct associations with the current situation.

Consider as an example a handball player whose team is poised to shoot on the opposition's goal. In handball each team consists of six field players and one goalkeeper. The teams each possess one goal (3 m wide × 2 m tall) at opposite ends of the playing field. There is a semicircle (6-m radius) extending from the goal in which only the goalkeeper is allowed to be; the remaining defenders are typically positioned around the semicircle. A player possessing the ball can hold it for up to 3 s or take three steps, before being required to dribble the ball, pass the ball to another player, or shoot at the goal.

If the primary playmaker has the ball in the middle of the field and is generating possible courses of action that player may adopt either a functional or a spatial strategy, which are typically used in tactical training in handball (Gerard, 1978; Mariot & Delerce, 2000; Zantop, 1986). For example, perhaps training and past game experience have led the playmaker to consider possible scoring opportunities before passing, or maybe the current situation suggests that retaining possession is important and thus passes should be considered first. These strategies would be labeled "functional" because they focus on the function or outcome of potential actions. Alternatively, maybe the defensive players on the left side of the field have been struggling during the current contest and so the playmaker first considers all possible options on that side of the field (a "spatial" strategy, defined by orientation rather than outcome). In each of these situations, different sequences of options may be generated, including "shoot left" and "shoot right" in the first case; "pass left" and "pass right" in the second case; and "shoot left" and "pass left" in the third case. Thus, according to our conceptual model, the same situation may produce different option-generation patterns depending on the strategy. These generated options would lead to the generation of additional, strongly connected (in memory) options. These association strengths can be acquired by experience or instruction such as described by Gerard (1978) and Mariot and Delerce (2000).

The elegant simplicity of the TTF heuristic is that it incorporates both the individual's domain expertise and the immediate context in determining how cognition proceeds. The heuristic is very much in the spirit of expert-based and naturalistic approaches to decision making (Lipshitz, Klein, Orasanu, & Salas, 2001; Salas & Klein, 2001). For example, Klein, Wolf, Militello, and Zsambok (1995) applied similar ideas to the generation of possible moves by experts and novices in chess. Assuming that chess players use a serial generation and evaluation of options (moves), they found that expert players only generated about three to five moves, and that the option-generation process systematically followed "move quality" in that qualitatively better moves (determined by expert rating) were generated first (see also de Groot, 1965). We have both option-generation and choice data of male handball players that further supports our model (Johnson & Raab, 2003). These athletes produced on average relatively few options, after which they chose mostly the option that came to mind first—hence the TTF heuristic's name. The fact that these initially generated options were also on average better than options produced after more extensive deliberation underscores the prescriptive merit of "taking the first" as well.

Open questions remain as a result of this previous work, which we address in the current research. First, the information search strategies are not yet well understood. The work of both Klein et al. (1995) and Johnson and Raab (2003) inferred these strategies mainly from outcome measures. Second, differences resulting from varying levels of expertise were not clear as a result of the previous work. We overcome both of these limitations in the current work, by additionally classifying strategies based on information search patterns measured by eye tracking, and collecting longitudinal data from groups with varying expertise.

Study Design and Predictions

We developed a longitudinal design in which three groups of participants with different levels of expertise were tested in four waves (labeled T1 to T4) occurring approximately every 6 months over a period of about 2 years from October 2004 to April 2006. All measures were gathered at one time from an individual participant within a wave and all participants were tested within a 2-week period for a given wave. The typical design with age-matched expertise versus novice groups was not pragmatic for our study, as true novices could not meaningfully complete the option-generation task. Furthermore, abundant evidence suggests that the lack of age matching is not critical to the current study, as chronological age is not a strong predictor of decision making skill in sports (for an overview, see McMorris, MacGillivary, Sproule, & Lomax, in press). For instance, McMorris et al. found strong effects of expertise but none for chronological age when predicting decision-making performance in soccer among nearly 300 young adults with a similar age range to our sample. We did collect a number of variables such as age and training age as well as a number of control variables such as amount of tactical training during the longitudinal study, tactical knowledge, and perceptual recognition.

A number of predictions can be made for the current study because it comprehensively includes the important components of search, generation, and choice. Where possible, the specific quantitative results of Johnson and Raab (2003) are used to estimate effect sizes for the current study, where Cohen's f was used for the multiple-group comparisons with $f = 0.25$ interpreted as a medium effect and $f = 0.40$ as a large effect. However, we also make novel predictions beyond the scope of the TTF model as originally formulated. These result from (a) our new assumptions about the influence of visual attention, and (b) our prediction of both cross-sectional (different training levels) and longitudinal (different waves) expertise effects. Predictions 1 to 3 are collapsed across expertise levels as we think they should be robust over different levels of expertise.

Predictions 1: Information Search (Gaze Characteristics)

We predict that participants will employ one of two gaze strategies (1a: Gaze strategy). The functional gaze strategy scans the full visual field and consequently produces a greater number of fixations of shorter duration. The spatial gaze strategy scans only part of the field and results therefore, in a smaller number of fixations with longer duration. The prediction was derived from our model assumption that gaze reflects the regions of interests based on a "top down," expertise-driven search strategy. We expect to find a positive relationship between gaze strategy and option-generation strategy—the sequence with which options would be verbalized is determined by the sequence of fixation areas in which these options are present (1b: Fixation order). Specifically, we predict that the initial option verbalized would most often be found in the area that was fixated first. The rationale is that, just as with perceptually salient stimuli, one's initial or intuitive solution to a problem will receive immediate attention when the scene is frozen. We also want to explore whether the fixation duration in a given region of the field correlates with the likelihood of choosing options located within the region (1c: Fixation duration). One rationale is that people may tend to confirm their initial option, as indicated by the confirmation bias in many cognitive tasks (Wason, 1960), by prolonging attention to the corresponding region. However, the alternative result also seems plausible—namely, that participants would generate an initial option in one region and then search primarily in other regions as a way of "ruling out" the remaining possible options in these regions. Either way, the results would enable us to further elaborate our model of the influence of gaze behavior on choices.

Predictions 2: Option Generation

We predict that the initial option generated would be better than the sequentially later-generated options (2a: Quality of generations), based on our model's success-driven option-generation process, described earlier. We seek to not only replicate this result from Johnson and Raab (2003), but to generalize it to different levels of expertise. We expect extensive use of the TTF heuristic, which suggests generation of only a small number—two to three—of options (2b: Number of generations), again replicating and generalizing the results of Johnson and Raab. Finding generations of more than one option would also indicate the presence of option generation as opposed to pure mapping of a situation and a single action, as proposed by alternative models (Phillips, Klein, & Sieck, 2004). We further predicted that a functional generation strategy would result in more options generated compared to a spatial strategy (2c: Generation strategy). The rationale is based on the fact that a spatial strategy would limit options to a specific region, whereas a functional strategy could generate options across the entire field, resulting in more options (cf. Johnson & Raab, 2003). Based on our previous results, we predicted a strong effect of strategy on the number of generated options ($f = 0.5$ in Johnson & Raab, 2003).

Predictions 3: Choice

We predict that participants will primarily choose the first option that comes to mind (3a: Take the first). We predict a replication of this result from Johnson and Raab (2003), with a TTF rate of about 60%. We predict that a functional generation strategy would result in lower quality of options compared to a spatial strategy (3b: Choice quality). The rationale is based on the fact that a spatial strategy is predicted to produce more options (2c) and an increase in generated options is predicted to result in lower quality (2a). We predicted a strong effect for this relationship, which also follows from Johnson and Raab ($f = 0.4$).

Predictions 4: Expertise

Perhaps the most important analyses are those that not only identify effects, such as those predicted above, but also examine how expertise plays a role in moderating these effects. In our previous investigations (e.g., Johnson & Raab, 2003), we had no specific predictions about expertise, although we alluded to connections between our work and the organization of expert memory (e.g., Chase & Ericsson, 1982; Ericsson & Kintsch, 1995). We predict expertise to reveal both cross-sectional differences and longitudinal changes. These analyses would allow us to discern both relatively short-term (over months) adaptations within individuals, as well as relatively long-term (over years) adaptations that occur as a function of domain-specific experience. Although a longitudinal investigation spanning only 2 years is likely to find only small changes in overt behaviors, the age groups in the current study are ideal for potentially witnessing changes during this period (McMorris, 1999). Longitudinal data reveals improvements intraindividually as well as differences interindividually and allows us to attribute these changes to expertise improvements or developmental differences in age.

4a: Cross-sectional effects. The TTF heuristic was derived in the tradition of simple heuristics that adaptively capitalize on domain-specific experience (Raab & Gigerenzer, 2005). Thus, we predicted that increasing expertise would correspond to increasing use of the TTF heuristic. This should be characterized by better and faster choices of experts compared to near-experts and nonexperts (see Williams et al., 2004). In addition, we predicted expertise-based differences in option-generation strategy such that experts would use a more spatial option-generation strategy compared to near-experts and nonexperts. This hypothesis is derived from our elaborated model and is neither predicted nor tested in previous literature. Furthermore, we predicted more spatial gaze strategies by experts as well, reflected by a higher number of fixations of longer duration in one area of the visual field compared to near-experts and nonexperts, in line with previous results (see Williams et al., 2004).

4b: Longitudinal effects. We predict that changes in all three decision-making components (search, generation, and choice) would occur across the longitudinal time span of the study. We predict that, across waves: The search rule would change from a more functional to a more spatial strategy; the initial option would be chosen more often (resulting also in better choices overall); and there will be reductions in the number of options generated, the number and duration of fixations, the time to generate the initial option, and the total option-generation time.

Method

Participants

A total of 90 participants agreed to take part in this study. However, only 69 took part in all four waves and consequently our analyses focus only on these individuals. The moderate drop-out rate of the near-expert and expert groups is based on transfers to different states and clubs in the course of this 2-year study, interrupting continuous data collection.

The 69 expert, near-expert, and nonexpert male and female handball players (for demographics and control variables see Table 1) were recruited from the state training center and clubs in north Germany. In Germany, the national organizing body in handball runs a club system that classifies teams based on both skill and age (in contrast to, say, Little League baseball/softball's solely age-based divisions in the United States). The participants in the current study were recruited specifically to represent three selected samples with differential expertise or skill. All participants were successful in their respective divisions: the experts were the National champions at their level, the near-experts were nationally competitive, and the nonexperts were from among the top local teams. All participants provided informed consent before participating in the study, which was carried out according to the ethical guidelines and with the approval of the University of Flensburg.

Table 1
Group Comparisons on Demographic and Control Variables at Baseline

	Experts (M: $n = 19$; F: $n = 10$)		Near-experts (M: $n = 13$; F: $n = 9$)		Nonexperts (M: $n = 8$; F: $n = 10$)				
	M	*SD*	*M*	*SD*	*M*	*SD*	*MSE*	*F*	*Cohen's f*
Age	17.14	(0.84)	15.23	(0.61)	13.98	(1.52)	84.33	42.85**	5.16
Years of training	9.45	(2.69)	8.55	(1.95)	7.10	(2.29)	261.16	4.57**	0.55
Tactical knowledge	3.00	(0.47)	3.12	(0.73)	3.04	(0.64)	90.49	0.24	—
Perceptual recognition	0.94	(0.10)	1.02	(0.11)	1.06	(0.19)	1.16	2.02	0.24
Tactical training	39.50	(5.85)	35.00	(2.05)	26.00	(1.89)	70.57	4.19	0.50

Note. Age and years of training are at Wave 1; tactical knowledge maximum is 5 points; perceptual recognition is in cm; tactical training is percentage of total training devoted to tactics, obtained with a chi-square analysis of different content between groups. $df = 68$. Effect size values are only displayed for $F > 1.0$. M = male; F = female.
* $p < .05$. ** $p < .01$.

Materials and Measurement

Decision-making video test. The video test procedure was adopted from Johnson and Raab (2003), except new sequences were employed. The video clips represent situations that were typical for the level of expertise of the current sample and adjusted to the calibration procedure of the eye tracking. The video test consisted of 15 clips of about 10 s each, taken of a near-elite team in training simulating competition-like situations (4 clips were used for warm-up). The clips were all filmed from a camera position 2 m above the ground from the line in the middle of the field such that all offensive and defensive players were visible. Each clip showed the development of a clearly identifiable attack situation, starting when the team sets up an attack and stopping (freeze-frame) when one of the attacking players in the middle of the field gains possession of the ball. The frozen frame was held for 6 s during which participants were asked to verbally generate options of the player in possession of the ball with the time of each option generated, position of each option generated, and number of options generated being digitally recorded. The selection of video clips was pilot tested with four validation procedures including expert (coaches) ratings of similarity to realistic situations and matching to specific defensive and offensive tactical components, and by an item analysis selecting items of different degrees of difficulty (see Johnson & Raab, 2003). In addition, another two coaches (from club teams within the highest level of expertise) selected the optimal point of freezing the attack situation to allow for multiple possible courses of action of varying quality (based on the development of the play and player alignment). The use of a frozen frame ensured a constant situation for option generation (i.e., that the first option would be generated in the same situation as the last option) and a consistent and reliable map of the ensuing eye-tracking data to the positions of the players. This type of task has been used earlier in research involving offensive situations in handball (Johnson & Raab, 2003), basketball (Raab & Johnson, 2004), and other sports (e.g., volleyball; Raab, 2003).

Eye tracking. A video-based head-mounted infrared eye tracker from BioMed Jena (2003, version 2.0) was used. Sampling rate was 25 frames per second with a spatial resolution precision of 0.01°, a spatial error of less than 0.8°, and a visual angle precision error less than 1°. The system was able to measure eye movements of about 29° horizontally and 19° vertically. Small fixations (< 120 ms) were filtered out because the link to cognitive processes seems valid only if the thresholds of the software filters in BioMed eye tracking systems were used (BioMed Jena, 2003, version 2.0; Williams & Davids, 1998). In addition, fixations were removed if they were outside a preset threshold in the *x* and/or *y* direction of the playing field, for instance, if the participant fixated on a side wall or the ceiling.

Gray adjustments of the individual eye were done by a histogram analysis to receive optimal solutions of offsets in detecting the pupil. Contrast and brightness of the eye camera were individually adjusted. Before calibration the regions of interest of the eye were fixed to enable optimal search for the pupil. Eight participants (all female) were required to remove eye makeup to facilitate pupil tracking. Calibration before the test used the software's 9-point monitor calibration, which was repeated until the system accepted the match. The 9-point calibration was checked before and after the warm-up trials.

Control Variables

Tactical knowledge. Domain-specific knowledge is built up over years of experience and influences the kind and number of options that are evaluated before a decision (McPherson & Kernodle, 2003). The tactical questionnaire is a test of knowledge for attack situations based on the training plan developed by the National coaches and used in the levels of play of our participants. We selected five different situations that represent typical knowledge that should be present in all expertise levels in our study. The multiple-choice questions were written and presented visually in a typical form familiar to the players. The test was previously used and validated to test tactical knowledge in handball (Johnson & Raab, 2003; Raab, 2003). Reliabilities in this study using Spearman–Brown coefficients based on the complete sample for the first wave showed a coefficient of $r = .31$, $p < .05$, $df = 68$. Validity of the same questionnaire as used in this study was previously investigated (Raab, 2003), and revealed a correlation between players knowledge and expert ratings of their knowledge of $r = .86$, $p < .01$, $df = 19$.

Perceptual memory recall. Williams and Ward (2003) used a pencil-and-paper test to measure the recalled position of particular players in soccer compared to their actual position in a previously displayed attack sequence. We used a perceptual memory recall test similar to this and adapted it to the handball situation. Specifically, after two practice trials we used six scenes (pictures of the frozen frame of the video test) in a computer-based version of the perceptual memory recall test. The six scenes represented the two main defensive and offensive structures used in the video test. Participants saw each scene for 5 s; each scene was then masked and participants had 2 min to reconstruct the position of the six attack and six defense players and the position of the ball on an empty handball half field with identical perspective and size as the scene presented (17-inch monitor with a handball half field of 270×202 pixels, height $\times$ width). The participants reconstructed the players' positions using symbols for attackers and defenders familiar to them from blackboard drawings used in their tactical training. They used a computer mouse to drag and drop the symbols from outside the represented field to positions on the field. The number of attacker and defender symbols to be dragged by the computer mouse matched exactly the number of attackers and defenders on the video clips. We overlaid the symbols on the images and calculated root mean square error between the foot position of these players and the middle of the overlaid symbol. Root mean square error averaged over all positions and scenes was the dependent variable of perceptual memory recall.

Training amount and content. One problem researchers face in running longitudinal studies is that there are a number of influencing factors that can systematically moderate the performance of individuals and groups. Therefore, we asked the participants' coaches to fill out a report in which they were asked to note injuries of players as well as an estimate of the percentage of training devoted to tactical issues (0 to 100%) before each test in the longitudinal study for the period between the tests. In addition, we asked them to mark on four scales (0 to 100%) the percentage of training that was spent on individual tactics, group tactics, defense tactics, and attack tactics. A simple analysis of amount and content of training was used to determine if there were differences between expertise groups or test intervals that needed to be taken into account in our interpretation of results.

Procedure

After providing informed consent, participants were seated and the eye-tracking apparatus was mounted and calibrated. Participants then read instructions on the screen about the video test. They were asked to verbalize the first option that came to mind when the scene was frozen (hereafter, "initial option"), then to generate other appropriate options (if any), and finally to pick from the list of options just generated the one they thought was best (hereafter, "final option"). Before and after each of the 15 video clips, a scene calibration was performed by the participants pressing a button when their eyes were fixated on a red circle on the screen. We included a one-point calibration also after the clips in case of the need for backward calibration in long option-generation periods; however this was not applied in data analysis due to reliable eye-tracking data from scene calibration before each scene. After the main task, participants were released from the eye-tracking apparatus to complete the supplemental tasks (e.g., tactical knowledge questionnaire and perceptual recognition test), after which they were debriefed. This entire procedure, including the supplemental tasks, was carried out in the same order for all four waves.

Data Analysis

All dependent variables were normally distributed after outlier reduction. In total 1.4% outliers were present, summed over all variables and expertise groups. There were no expertise groups or variables that systematically produced a large number of outliers. Outliers were reduced by replacing all values per expertise group by gender that were higher or lower, such that all values per expertise group by gender that were higher or lower than two standard deviations from the group mean were replaced by the value of two standard deviations.

Verbal option-generation data. Digitally recorded options were coded by two raters ($r = .88$, $p < .01$, $df = 68$). We used a master list that numerically coded all free response options of the participants based on a list from Johnson and Raab (2003) that contained 107 different options. The high number of options in this task is a result of the variety of plausible moves including slight variations (e.g., a bounce pass to the left wing vs. a lob pass to the left wing vs. a pass to the left wing after a fake pass to the right). Also, the serial position and time for initial option generated as well as the generation time (1-ms precision) were matched with each option, for each scene, for each participant. Time from stopping the scene to initial option generated was measured by the software. A tone signal was played on the soundtrack of the video clip at the moment the scene stopped and the software coded the time when the digital recording measured the beginning of the verbal response. Option-generation time was measured from the start of the scene to the last option generated. The options were then coded as appropriate or not appropriate based on two coaches' rating of these outcomes, regardless of the specific method (e.g., if a pass to the left wing was rated as appropriate, any option with this outcome was coded as appropriate). Only the most appropriate option, as determined across coaches, was classified as "correct" in computing decision quality. The decision quality for a participant's initial options was calculated by determining how many trials the initial option corresponded to the coaches' "correct" option (and likewise for options in other generation positions and the final option). The mean (across scenes) number of options generated for each participant was also calculated. Reliability scores for decision and gaze measures were calculated using the split-half test (see Table 2).

The sequence of options generated for each scene was analyzed and assigned to one of the two option-generation strategies (spatial or functional). If a sequence of generated options included primarily just one third of the field (left, middle, or right), regardless of whether the options resulted in passes or shots on goal, this sequence was classified as spatial. Sequences were attributed to a functional strategy if only passes or only shots on goal were generated, regardless of the position of the options on the playing field. Participants were labeled as functional or spatial if one of these option-generation strategies was used in the majority of video clips. Option-generation sequences with fewer than three options were excluded because classification was not easy to achieve with such a small set—for example, a sequence of "pass to the left wing player" and "pass to the left halfback" could be classified as either functional (only passes are generated) or spatial (only options on the left side are generated). In all other calculations of choice data all options were inserted into the calculations. In addition, option-generation sequences that could not be easily attributed to either strategy directly were coded as ambiguous. The rationale for separating the strategies into functional and spatial was that they represent the two main tactical training forms and reflect the empirical data from Johnson and Raab (2003).

Gaze characteristics. Eye-tracking data were collected from the start of the video clip until the first verbal utterance (initial option generation). We determined fixation using three equally sized regions of interest, defined vertically: the left field and right field (each containing two teammates) and the center field (containing the goal and one teammate). We calculated for each scene the number of fixations and mean fixation duration for each participant to each region. In addition, we categorized the sequence of fixations to one of many different gaze strategies, which can be summarized by the degree of spatial attention. Attention to

Table 2
Reliabilities of Decision and Gaze Measures

	T1	T2	T3	T4
Gaze classification	.68	.75	.83	.64
Fixation duration	.74	.57	.53	.73
No. of fixations	.82	.84	.86	.93
Option classification	.96	.92	.89	.92
Quality: First option	.49	.89	.69	.94
Quality: Final option	.66	.68	.68	.51
No. of options	.77	.84	.83	.86
Generation time: First option	.92	.89	.87	.99
Generation time: Total	.86	.75	.86	.90

Note. $N = 69$. Values represent Spearman-Brown coefficients based on the complete sample (coefficients by group result in comparable reliability scores and can be requested on demand). Gaze and option classification is based on two independent raters of all trials for each of the four waves. T1 to T4 represent the waves in the longitudinal study.

the entire visual field could be indicative of a functional gaze strategy, whereas attention to only a single (left, middle, or right) region suggests a spatial gaze strategy. Intermediate spatial attention patterns (e.g., focusing back and forth between two regions) are not as easily classifiable and are labeled as ambiguous. Note that the classification based on gaze characteristics is determined independently from the classification using verbal option generation.

Results

Significance criteria of $p < .05$ and $p < .01$ were established for all analyses. Prior to testing the main hypotheses, we examined the potential moderating effects of tactical knowledge and tactical training. There were no statistically significant moderating influences from these factors (see Table 1). Age typically does correlate with expertise; in this sample the correlation is $r = .75$, $p < .01$, $df = 68$. However, in our design we cannot exclude the potential confound of age and therefore we will discuss the expertise effects bearing in mind this potential confound.

For cross-sectional comparisons we calculated a 3 × 2 × 2 (Expertise × Generation Strategy × Gaze Strategy) multivariate analysis of variance (MANOVA) including as dependent variables: quality of the initial option, quality of the final option, number of generated options, decision time, generation time, dynamic inconsistency, number of fixations, and fixation duration. Expertise, generation strategy, and gaze strategy are between-participants factors.

Information Search (Gaze Characteristics)

Prediction 1a. We predicted that participants would employ either a functional or a spatial search strategy. Therefore we classified participants' search strategies, as described earlier, based on the amount of gaze focus on different regions. Of those participants that could be classified (i.e., nonambiguous classifications), the spatial strategy was employed by about 51% of experts, about 41% of near-experts, and about 55% of nonexperts. We explored whether gaze-strategy classification changed over the four waves and found that 48 of 69 participants exhibited exactly one change, and the remaining 21 participants used a consistent gaze strategy across all four waves. In 11 cases the change was from a spatial to a functional gaze strategy, and in 37 participants this change was from a functional to a spatial strategy, which we would expect to be the better transition. The proportion of participants switching strategies was significant, $\chi^2(1, N = 69) = 6.50$, $p < .05$, $f = 1.02$.

For a direct comparison we classified independent strategies based on option-generation sequences and found that a spatial strategy was employed by about 61% of experts, about 36% of near-experts, and about 59% of nonexperts. In a contingency table we found that option-generation strategies and gaze behavior corresponded such that spatial gaze and spatial option-generation occurred together 22 times; spatial gaze and functional option-generation 14 times; functional gaze and functional option generation 21 times; functional gaze and spatial option generation 12 times. A chi-square test revealed that this classification pattern was significantly different than an equal distribution based on marginal frequencies (i.e., a test of independence among classifications), $\chi^2(1, N = 69) = 4.22$, $p < .05$, $f = 0.83$.

Prediction 1b. We predicted that participants would fixate first on the region in which the intuitive and, most often, best option was located. If fixations begin randomly, we would expect within 11 scenes to have about 3.6 first gazes for each third of the display. We found that in about half of the scenes the first gaze was identical to the position of the initial option generated, which is significantly different from chance (for data in each wave see Table 4). We ran a similar analysis relating the final gaze position to the final choice, but this result was not significantly different from chance in any wave.

Prediction 1c. We assumed that fixation duration would be different in the area of fixation for high compared to low quality choices. Therefore, we analyzed if participants' fixation duration was longer in those areas that prompted them to name their initial option. We expected that longer fixation on a particular area would be indicative of participants choosing a good option in this area. We found as predicted that fixation duration was longest in the areas with which the initial option was associated in all four waves, with high effect scores (see Table 4).

Option Generation

Prediction 2a. We predicted that the initial option generated would be better than later generated options. An analysis of variance conducted with position of options as the within-participant factor and expertise group as the between-participants factor revealed significantly better quality of the initial option, compared to subsequent options, for each of the four waves (see Table 4). In addition we analyzed, for cases in which the initial and final options diverged, if the initial option generated was better than the final option. There was evidence for this in three of four waves. This indicates that participants were quite effective in picking a very appropriate option intuitively, which further deliberation was unable to improve.

Prediction 2b. We predicted, based on the results of Johnson and Raab (2003), that participants would generate only about two to three options and this was the case on average across all three expertise groups (T1: $M = 3.33$, $SD = 1.23$; T2: $M = 2.32$, $SD = .58$; T3: $M = 2.27$, $SD = .59$, T4: $M = 2.27$, $SD = .51$). Determining the significance of this result is quite difficult, as it is not informative to compare the number of generated options with the possible number of generations (over 100). Using a 95% confidence interval we found that option generation averaged over expertise groups was between 2.0 and 2.7 (T1), 1.2 and 1.5 (T2), 1.1 and 1.4 (T3), and 1.1 and 1.4 (T4). We also decided to simply check whether the number of options generated in this study differed from that reported in our previous study (Johnson & Raab, 2003) and thus ran a *t* test for each wave separately with the test value drawn from the previous study (2.5 options generated). This analysis revealed significantly more options generated in the first wave and significantly fewer options generated in all subsequent waves (see Tables 3 and 4).

Prediction 2c. We predicted option generation by functional search would result in more options being generated than spatial option generation. Most of the participants could be assigned to either a spatial or a functional strategy of option generation if more than two options were considered. The MANOVA (see Table 5) revealed no significant differences due to generation strategy, and

Table 3
Means and Standard Deviations (in Parentheses) by Expertise and Wave for All Dependent Variables

	Experts				Near-experts				Nonexperts			
	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4
Information search												
No. of fixations	16.73 (4.53)	9.54 (3.53)	5.27 (2.38)	6.16 (2.54)	18.66 (7.84)	11.34 (4.53)	4.97 (2.43)	5.78 (1.99)	17.70 (5.83)	10.48 (5.44)	8.23 (3.46)	5.55 (2.25)
Duration of fixations	5.67 (0.53)	4.91 (1.31)	2.61 (0.96)	5.76 (1.17)	5.54 (0.52)	4.85 (0.49)	2.32 (0.63)	2.60 (0.36)	5.65 (0.58)	4.66 (1.77)	3.71 (1.38)	2.09 (1.48)
Option generation												
Initial option quality	5.05 (1.35)	3.23 (1.42)	3.33 (1.49)	3.63 (1.15)	3.56 (1.70)	2.58 (1.38)	3.25 (1.52)	3.00 (1.56)	3.63 (1.30)	2.56 (1.15)	2.39 (1.14)	2.79 (.98)
Second option quality	1.00 (0.71)	0.88 (1.05)	0.92 (0.67)	0.63 (0.72)	1.44 (1.19)	1.05 (0.91)	1.00 (0.81)	0.30 (0.48)	1.42 (1.12)	0.96 (0.73)	1.45 (1.06)	0.50 (0.75)
Third option quality	0.56 (0.68)	0.29 (0.47)	0.50 (0.79)	0.63 (0.71)	0.55 (0.70)	0.42 (0.69)	0.44 (0.63)	0.30 (0.48)	0.55 (0.68)	0.52 (0.71)	0.32 (0.56)	0.50 (0.75)
No. of options	2.95 (0.98)	2.25 (0.62)	2.28 (0.52)	2.39 (0.56)	3.41 (1.27)	2.34 (0.59)	2.39 (0.63)	2.25 (0.55)	3.44 (1.30)	2.36 (0.55)	2.17 (0.58)	2.13 (0.42)
Generation time	24.07 (4.12)	12.86 (3.01)	12.24 (4.34)	15.96 (4.05)	24.47 (4.73)	14.83 (3.36)	14.88 (4.79)	15.53 (2.24)	23.19 (5.79)	14.72 (5.03)	11.78 (4.46)	14.96 (5.06)
Choice												
Final option quality	5.22 (1.55)	3.58 (1.41)	4.00 (1.34)	4.75 (1.23)	3.44 (1.13)	3.21 (1.23)	3.63 (1.45)	3.21 (1.23)	3.63 (1.52)	3.16 (1.14)	2.89 (1.12)	3.20 (1.31)
Decision time	3.75 (1.28)	3.49 (0.63)	2.95 (0.71)	3.02 (0.78)	3.98 (0.89)	4.05 (1.14)	3.36 (1.22)	3.12 (1.09)	4.92 (2.46)	4.25 (1.67)	3.62 (1.48)	3.38 (1.60)
Dynamic inconsistency	3.63 (2.41)	2.33 (1.97)	3.16 (1.89)	3.89 (1.74)	5.14 (3.83)	3.47 (2.34)	3.79 (2.39)	2.70 (1.76)	3.43 (1.30)	3.04 (1.71)	3.52 (2.56)	2.38 (1.48)

Note. $N = 69$. T1 to T4 = the four waves; first, second, third, and best option = number of correct choices based on expert rating; no. of options = options other than the best option; decision time = time in seconds between freezing of the scene and first option generated; generation time = time in seconds between start of the scene and final choice; dynamic inconsistency = the number of trials in which first and best choice are not identical; no. of fixations = the number of fixations averaged over test items in seconds; duration of fixations is in mean time of fixations in seconds.

Table 4
Analysis of Variance for Predictions 1 to 3

Wave	*MSE*	*F* or *t*	Cohen's *f*
Prediction 1b[a]			
T1		4.29**	0.54
T2		5.98**	0.75
T3		7.34**	0.93
T4		2.57*	0.32
Prediction 1c[a]			
T1		5.73**	0.72
T2		7.87**	0.99
T3		2.24*	0.29
T4		6.49**	0.82
Prediction 2a[b]			
T1	97.92	22.91**	1.06
T2	134.95	24.09**	0.93
T3	109.41	20.15**	0.66
T4	108.76	17.35**	1.00
Prediction 2b[a]			
T1		13.83**	1.74
T2		17.97**	2.26
T3		15.41**	1.94
T4		16.10**	2.03
Prediction 3a[a]			
T1		10.51**	1.32
T2		11.90**	1.49
T3		10.55**	1.33
T4		10.21**	1.29
Prediction 3b[b]			
T1	126.45	0.71	
T2	92.68[c]	10.80**	0.43
T3	89.18	10.07**	0.47
T4	81.77	10.65**	0.54

Note. Prediction 1a is in the text. Predictions 2c and 3a as well as predictions of expertise (4) are analyzed by means of the MANOVA (see Table 5). Number of generations was tested against a value of 1. Effect size values are only displayed for $F > 1.0$. Prediction 1b = first gaze fixation and initial option generated; Prediction 1c = fixation duration per area and choice in this area; Prediction 2a = quality of generations; Prediction 2b = number of generations; Prediction 3a = Take The First; Prediction 3b = choice quality.
[a] Reported statistic is t, $df = 61$. [b] F is reported, $df = 63$. [c] $df = 61$, due to two missing data entries.
* $p < .05$. ** $p < .01$.

thus we cannot conclude that different generation strategies produced different numbers of options.

Deliberation and Choice

Prediction 3a. Those not exhibiting the TTF heuristic instead show dynamic inconsistency, defined as a disparity between one's initial and final choices. Considering the dynamic inconsistency rate of 40% found by Johnson and Raab (2003), we would expect this phenomenon to appear on an average of about 4.4 of 11 trials; this test value was used in a *t* test. Using a 95% confidence interval we found that the number of inconsistent trials averaged over expertise groups was between 3.8 and 5.6 (T1), 2.5 and 3.5 (T2), 2.8 and 4.2 (T3), and 2.5 and 3.7 (T4). We found a higher dynamic inconsistency rate only for the first wave of the near-expert group, whereas all other groups and waves showed lower dynamic inconsistency. Collapsing across groups, a significantly lower overall rate was found in waves T2 to T4 (Tables 3 and 4).

Prediction 3b. Option generation by functional search should result in options with lower choice quality compared to spatial option generation. The MANOVA (see Table 5) revealed no significant difference for type of option generation, and thus we cannot conclude that there were differences across strategies in the quality of options generated.

Prediction 4: Expertise

Prediction 4a. We predicted that expertise would have an effect on the three decision-making components of search, generation, and choice. A MANOVA on the T1 data showed a strong effect of expertise, as predicted (see Table 5). A detailed inspection of the variables suggests this result stems primarily from expertise-based differences in quality of both initial and final options, although differences in number of generated options (experts generating fewer) and number of fixations (experts fixating less) are also apparent. However, dynamic inconsistency was a less prominent difference between levels of expertise, contrary to our predictions. All three two-way interactions were significant; however without specific predictions we will refrain from post hoc elaboration on these results.

Prediction 4b. For the longitudinal analysis we used wave as a within-participant factor with four levels and tested for effects on each dependent variable (see Table 6). The main changes appear to be in the gaze characteristics in which the number and duration of fixations decreased over the waves, with a particularly large effect ($f = 2.38$) found for duration. Generation time also significantly decreased across waves and represented a large effect ($f = 1.06$). No other differences were statistically significant, contrary to our predictions.

Discussion

The goal of the current study was to examine information search, option generation, and choice, as well as the interaction of these components, among athletes of varying levels of expertise. We were especially interested in the extent to which different search strategies (inferred from gaze characteristics) can predict subsequent deliberation (option generation and choice). We ap-

Table 5
Multivariate Analyses of Variance of Main Effects of Factors Expertise, Option Generation, and Gaze Behavior

Source	*F* (44)	Cohen's *f*
Expertise (E)	2.59**	1.09
Option generation (O)	0.58	
Gaze (G)	1.78	0.90
E × O	2.43**	1.05
E × G	2.11*	0.98
O × G	3.43**	1.25
E × O × G	0.98	
Error	(75.34)	

Note. Option generation represents a two-level (spatial and functional) factor classified by the verbal option-generation data. Gaze represents a two-level (spatial and functional) factor classified by gaze behavior (eye-tracking data). Value enclosed in parentheses represents mean square error. Effect size values are only displayed for $F > 1.0$.
* $p < .05$. ** $p < .01$.

Table 6
Analyses of Variance for Longitudinal Data

Variable	*MSE*	*F* (68)	Cohen's *f*
Information search			
No. of fixations	232.82	11.54**	0.85
Duration of fixations	45.73	26.37**	2.38
Option generation			
Initial option	73.88	1.15	0.26
No. of options	37.91	2.82	0.70
Generation time	738.00	22.36**	1.06
Choice			
Final option	79.51	0.63	
Decision time	60.86	3.15	0.75
Dynamic inconsistency	228.67	0.89	

Note. Effect size values are only displayed for $F > 1.0$.
* $p < .05$. ** $p < .01$.

proached this problem from a computational perspective that focused on the component processes involved in seemingly complex cognition. Our previous research (Johnson & Raab, 2003) produced some counterintuitive results, such as that generating fewer possible courses of action results in better choices, and that both option generation and final choice are strategy dependent. However, a direct measure of strategy (information search) was absent in the previous work, where strategy was instead inferred from the sequence of generated options. Furthermore, a formal comparison based on expertise was absent from our previous work, and other researchers had produced equivocal results. The current study fills some of these gaps by examining expertise-based differences in option generation and choice, and how these differences are affected by gaze strategies or information search.

We found support for the use of the Take-The-First heuristic (Johnson & Raab, 2003) for option generation of the participants of different expertise levels. Whereas our previous research showed effect sizes between about $f = 0.4$ and 0.5, we now found stronger effects for the quality of generations (ranging across the four waves from about 0.7 to 1.1), number of generations (ranging across the four waves from about 1.7 to 2.3), and use of TTF (ranging across the four waves from about 1.3 to 1.5). We attribute these large effects to the greater expertise of our participants, compared to those studied by Johnson and Raab. Participants generated only a few options, and quite often they chose the first option that came to mind. The current research illustrates that this heuristic seems quite stable across various levels of expertise and gender. Only about three options were generated on average, one more than found in a previous study (Johnson & Raab, 2003). The low number of generations was not likely due to a ceiling effect, or limitation of the number of possible actions in a given situation because many different possible courses of action were plausible in each situation (scene), and different participants identified varying options for a given scene. Because the sport and the attack situations in this study are the same as those in previous work in terms of complexity, the increase in mean number of generations could be due to greater variability in expertise of the participants in the current study. We predicted interactions between gaze classifications and option-generation strategies as well as interactions between each of these factors and expertise. We found moderate to large effects on choice and gaze variables indicating, in principle, our model is accurate.

We suggest two possible explanations for these results. First, even at the top level of these age groups, it is likely to be advantageous in high-pressure, real game situations to generate a small number of options, rather than inundating oneself with possible courses of action and becoming paralyzed with indecision. Second, there could have been an artificial task effect because participants may have felt compelled to generate options even though they had already produced what they truly felt was the correct alternative. However, this latter explanation was not supported by the debriefing session. These possibilities are only speculative. We think that all the groups in this study had considerable expertise that may have allowed them via training to adapt to fast decision making, and therefore no differences appeared. We will outline briefly below how further research may clarify this issue.

Participants may have indeed generated options due to experiment demands; that they frequently selected as their final choice their first generated option may be evidence that the subsequently generated options were "artificial" and not real candidates. However, even the highest expertise group selected the first-generated option only 60% of the time, indicating that there was additional processing and active comparison taking place. Similarly, Johnson and Raab (2003) found their participants used the TTF heuristic about 60% of the time. This is in contrast to findings in other domains, such as firefighting, where it is assumed that there is almost no option generation at all because only the one option that comes to mind first is immediately chosen (Klein, 1989).

The option-generation process found in this and previous research does not seem to match models that propose a random option-generation process (Chan & Courtney, 1998) or support the conjecture that there is no option generation per se in experts (Klein, 1989). Furthermore, models that do acknowledge option generation may need to be expanded in more detail (Klein et al., 1995; Phillips et al., 2004, p. 304, Figure 15.1). Phillips et al. proposed that the experience of a changing context requires generating and choosing among 1 to $(1 + n)$ actions, but did not specify important details such as how the strategy by which relevant cues influence the option-generation process. In addition, the structural differentiation of spatial or functional strategies and the resulting difference in amount, sequential order, and type of options generated are not predictable by Phillips et al.

Furthermore, there were no significant expertise-based differences in the tendency to choose the initial option. We refrain from further interpretation of the failure to reject the null hypothesis and leave that open for further investigations. In situations such as the current sports task in which good options come to mind first, it may in fact be prescriptive to "take the first" rather than exhaustively generate options. The quality of the generated options decreased with serial position, suggesting that higher quality options occurred earlier in the generation sequence (which replicates Johnson & Raab, 2003).

There were many qualitative and quantitative effects of expertise on performance. For example, as expected, the quality of initially generated options and finally selected options was influenced by expertise, a result stemming largely from the higher quality options generated and chosen by the expert group. There were also expertise-based differences in the speed with which the initial "intuitive" option was generated. This is interesting especially because the total number of generated options was about the same between groups but experts responded faster. In sum, these results suggest that greater expertise did not increase the number of possible actions generated or the time spent doing so, but it did increase the quality of those that were initially generated as well as the speed with which they came to mind. We believe that the locus of the expertise effect is in the accessibility and quality of an initial solution to a problem, not in subsequent deliberation (cf. the "leverage point" of Klein et al., 1995).

There were also significant expertise-based differences in visual information search—cautiously interpreted as a crude measure of attention—as measured by gaze fixation. Previously, researchers (e.g., Williams et al., 1999, 2002) have been inconclusive as to whether expertise leads to fewer or more fixations. If we analyze the global pattern then we find partial evidence for the often reported trend that experts require fewer fixations.

The tendency for experts to adopt a more spatially oriented strategy, as evidenced by both eye tracking and verbal protocol, seems to be the link that explains expertise-based differences on the current task. Specifically, we found that gaze strategy had no effect on the number of options generated, but that a spatial gaze strategy led to better initial generations and final choices. This finding alone can account for the expertise-based outcome differences of the current study, assuming that experts use extensively spatial strategies and near-experts use primarily functional strategies. More interesting, differences in the outcome measures were only evident using the gaze-based strategy classification, and not the generation-based strategy classification. This lack of apparent differences when using the generation-based strategy classification highlights the utility of the methods employed in the current study. Furthermore, it may help explain in part differences to the corresponding results of Johnson and Raab (2003) concerning expertise-based differences in strategy use, which used only the generation-based strategy classification. It is important to note that the results attributed to expertise potentially are confounded by age. As mentioned in the Introduction, evidence suggests that training age explains more variance than biological age for sport choices like those in the current study (e.g., McMorris, 1999; McMorris et al., in press). However, this does not dismiss a high correlation between expertise and biological age in our current study. Therefore we cannot attribute the reported effects solely to expertise. We believe that further research designed specifically to address this topic seems warranted.

The results of the present study, and the discussion above, show that in the process of answering some of our questions, we have generated several new areas of inquiry that should be addressed in the future. First, cross-sectional research on different expertise groups can hardly answer how people develop a TTF-type heuristic. It seems warranted to conduct further longitudinal studies to advance our understanding of how experts exhibit such behavior in the first place. The current study is in fact the first in a planned series meant to address this question. For instance, our conceptual model assumes that options are generated by association strength and similarity between options. However, the current study cannot easily distinguish and weigh these processes in option generation. Future researchers need to employ continuous ratings for association strength between a situation and an option as well as similarity ratings between options and potentially new study designs to distinguish and weigh these processes.

Second, more light needs to be shed on the parameters that influence the final choice. For instance, the duration of the assessment phase (and number of options generated therein) may influence the relative contribution of primacy and recency effects. This factor may be manifest in the degree to which the initial option is retained as a final choice as well as in the tendency for a later (in the generation sequence) option to be selected.

Third, there have been only a few intervention studies examining to what extent gaze strategies can be trained and decision making in the field improved (see Abernethy, 1991; Farrow & Abernethy, 2002; Jackson, 2003; Raab, 2003, 2007; Vickers, Livingston, Umeris-Bohnert, & Holden, 1999; Vickers, Reeves, Chambers, & Martell, 2004). In addition to studying the cognitive processing that occurs during option generation, it is imperative to consider decision making from a learning perspective. Unfortunately, this approach has been largely absent, which is particularly regrettable given the high level of benefits (i.e., explanatory power) that can be expected.

Fourth, it is also possible to use other research to more precisely define aspects of the computational model taken for granted in the current and past applications. For example, a great deal of work has examined the nature of experts' memory and differences in memory structure and storage/recall processes among experts and nonexperts (Ericsson, 2003). This research can be used to specify the assumed associations and association strengths in the computational model. It will also be informative to determine additional factors that interact with expertise in determining the tendency to adhere to the TTF heuristic, such as working memory (Beilock & Carr, 2001) or perceptual expertise (Williams & Ward, 2003). For example, perhaps the superior "chunking" ability of experts that integrates knowledge effectively expands their working memory, which in turn influences the degree to which preference for initial options is retained despite generation of subsequent options. Perceptual differences between experts and nonexperts would, at the most fundamental level, change the definition (or at least representation) of what constitutes "an option" and may also affect saliency cues that drive gaze behavior. Connecting the current model to these lines of research would be extremely fruitful and contribute to a more comprehensive understanding of expert cognition.

Fifth, although sports tasks provide a useful test bed for our model, it is important to generalize the findings of the current study beyond the sports domain. We have no reason to believe that the current results are specific to the sports domain, as it possesses important surface characteristics such as limited time, high pressure, and dynamic information that describe many real-world settings. Our model could be easily applied to situations studied within the naturalistic decision making approach, such as those faced by military commanders, firefighters, radar operators, and pilots. Furthermore, our model relates to properties of option generation reported in other applied domains such as parking and selecting living spaces (Gettys, Pliske, Manning, & Casey, 1987) and creative brainstorming (Kramer, Kuo, & Dailey, 1997). Other popular applied domains in decision research, such as the medical domain, also involve problems that can be cast in terms of strategy use (e.g., prevention vs. treatment), information search (e.g., diagnostic testing), option generation (e.g., possible treatments), and action selection. Future work could explore the value of our model in these other domains.

Finally, we believe that a theory of expertise has yet to be defined (or elaborated, Ericsson, 2003). The heuristic approach discussed here may be one effective way to develop descriptions of the decision making process of experts, test these processes and understand them as well as develop interventions that would enhance the steps of the expertise approach outlined so far.

References

Abernethy, B. (1991). Visual search strategies and decision-making in sport. *International Journal of Sport Psychology, 22,* 189–210.

Bard, C., & Fleury, M. (1976). Analysis of visual search activity during sport problem situations. *Journal of Human Movement Studies, 3,* 214–222.

Bar-Eli, M., & Raab, M. (Eds.). (2006). Judgment and decision making in sport and exercise: Rediscovery and new visions [Special Issue]. *Psychology of Sport & Exercise, 7*(6).

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130,* 701–725.

BioMed Jena GmbH. (2003). *Handbuch für das biovision eye-tracking* [Handbook of biovision eye-tracking] Version 2.0. Jena, Germany: Jena Selbstverlag.

Brisson, T. (2003). Player's perspective: Therese Brisson. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 216–218). Champaign, IL: Human Kinetics.

Chan, H. S., & Courtney, A. J. (1998). Revising and validating the random search model for competitive search. *Perceptual and Motor Skills, 87,* 251–260.

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 1–58). San Diego, CA: Academic.

de Groot, A. D. (1965). *Thought and choice in chess.* The Hague, The Netherlands: Mouton.

Ericsson, K. A. (2003). How the expert performance approach differs from traditional approaches to expertise in sport: In search of a shared theoretical framework for studying expert performance. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 371–402). Champaign, IL: Human Kinetics.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102,* 211–245.

Farrow, D., & Abernethy, B. (2002). Can anticipatory skills be learned through implicit video based perceptual training? *Journal of Sports Sciences, 20,* 471–485.

Gerard, K. (1978). Wing play in team handball. *Team Handball News, 2,* 11–13.

Gettys, C. F., Pliske, R. M., Manning, C., & Casey, J. T. (1987). An evaluation of human act generation performance. *Organizational Behavior and Human Decision Processes, 39,* 23–51.

Helsen, W., & Pauwels, J. M. (1993). The relationship between expertise and visual information processing in sport. In J. L. Starkes & F. Allard (Eds.), *Cognitive issues in motor expertise* (pp. 109–134). Amsterdam: North-Holland.

Jackson, R. C. (2003). Evaluating the evidence for implicit perceptual learning: A reanalysis of Farrow and Abernethy (2002). *Journal of Sports Sciences, 21,* 503–509.

Johnson, J. G., & Raab, M. (2003). Take The First: Option generation and resulting choices. *Organizational Behavior and Human Decision Processes, 91,* 215–229.

Klein, G. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), *Advances in man-machine system research* (Vol. 5, pp. 47–92). Greenwich, CT: JAI.

Klein, G., Wolf, S., Militello, L., & Zsambok, C. (1995). Characteristics of skilled option generation in chess. *Organizational Behavior and Human Decision Processes, 62,* 63–69.

Kramer, M. W., Kuo, C. L., & Dailey, J. C. (1997). The impact of brainstorming techniques on subsequent group processes: Beyond generating ideas. *Small Group Research, 28,* 218–242.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making, 14,* 331–352.

Mariot, J., & Delerce, S. (2000). Sports collectives: Nouvelle approache didactique functionnelle [Team sports: New didactic and functional approach]. *Education physique et Sport, 50,* 61–66.

Martell, S. G., & Vickers, J. N. (2004). Gaze characteristics of elite and near-elite athletes in ice-hockey defensive tactics. *Human Movement Science, 22,* 689–712.

McMorris, T. (1999). Cognitive development and the acquisition of decision-making skills. *International Journal of Sport Psychology, 30,* 151–172.

McMorris, T., MacGillivary, W. W., Sproule, J. & Lomax, J. (in press). Cognitive development and performance of 11, 13 and 15 year olds on a soccer-specific test of decision making. *International Journal of Sport and Exercise Psychology.*

McPherson, S. L., & Kernodle, M. W. (2003). Tactics, the neglected attribute of expertise: Problem representations and performance skills in tennis. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 137–168). Champaign, IL: Human Kinetics.

Phillips, J. K., Klein, G., & Sieck, W. R. (2004). Expertise in judgment and decision making: A case for training intuitive decision making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (p. 297–315). Malden, MA: Blackwell.

Raab, M. (2003). Decision making in sports: Implicit and explicit learning is affected by complexity of situation. *International Journal of Sport and Exercise Psychology, 1,* 310–337.

Raab, M. (2007). Think SMART, not hard—Teaching decision making in sports from an adaptive cognition perspective. *Physical Education and Sport Pedagogy, 12,* 1–22.

Raab, M., & Gigerenzer, G. (2005). Intelligence as smart heuristics. In R. J. Sternberg, J. Davidson, & J. Pretz (Eds.), *Cognition and intelligence* (p. 188–207). Cambridge, England: Cambridge University Press.

Raab, M., & Johnson, J. G. (2004). Individual differences of action-orientation for risk-taking in sports. *Research Quarterly and Exercise Sport, 75,* 326–336.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Salas, E., & Klein, G. (Eds.). (2001). *Linking expertise and naturalistic decision making.* Mahwah, NJ: Erlbaum.

Savelsbergh, G. J. P., Williams, A. M., van der Kamp, J., & Ward, P. (2002). Visual search, anticipation and expertise in soccer goalkeepers. *Journal of Sports Sciences, 20,* 279–287.

van Gog, T., Paas, F., & van Merrienboer, J. J. G. (2005). Uncovering expertise-related differences in troubleshooting performance: Combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology, 19,* 205–221.

Vickers, J. N., Livingston, L., Umeris-Bohnert, S., & Holden, D. (1999). Decision training: The effects of complex instruction, variable practice and reduced delayed feedback on the acquisition and transfer of a complex motor skill. *Journal of Sports Sciences, 17,* 357–367.

Vickers, J. N., Reeves, M. A., Chambers, K. L., & Martell, S. G. (2004). Decision training: Cognitive strategies for enhancing motor performance. In A. M. Williams & N. J. Hodges (Eds.), *Skill acquisition in sports* (pp. 103–120). London: Routledge.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12,* 129–140.

Williams, A. M., & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. *Research Quarterly for Exercise and Sport, 69,* 111–128.

Williams, A. M., Davids, K., Burwitz, L., & Williams, J. G. (1992). Perception and action in sport. *Journal of Human Movement Studies, 22,* 147–204.

Williams, A. M., Davids, K., Burwitz, L., & Williams, J. G. (1994). Visual search strategies of experienced and inexperienced soccer players. *Research Quarterly for Exercise and Sport, 65,* 127–135.

Williams, A. M., Davids, K., & Williams, J. G. (1999). *Visual perception & action in sport.* London: Spon.

Williams, A. M., Janelle, C. M., & Davids, K. (2004). Constraints on the search for visual information in sport. *International Journal of Sport and Exercise Psychology, 2,* 301–318.

Williams, A. M., & Ward, P. (2003). Perceptual expertise: Development in sport. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 219–250). Champaign, IL: Human Kinetics.

Williams, A. M., Ward, P., Knowles, J. M., & Smeeton, N. J. (2002). Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology: Applied, 8,* 259–270.

Zantop, B. (1986). Das taktische Verhalten des Angreifers in der Überzahlsituation 2:1 [Offensive tactics in the advantage situation of two on one]. *Sportunterricht, 35,* 187–193.

Received November 30, 2006
Revision received June 24, 2007
Accepted June 29, 2007 ■

Ad Hoc Reviewers – Journal of Experimental Psychology: Applied – Special Issue (September 2007)

Bruce Abernethy	Frank Durso	Nicola Hodges	Georgios Rigas
Michael Bar-Eli	David Eccles	Sean Horton	Remy Rikers
Thomas Busey	Damian Farrow	Gary Klein	Karol Ross
Richard A. Carlson	Paul Feltovich	Andreas Lehmann	Pertti Saariluoma
Thomas H. Carr	Michel Ferrari	Robert Malina	Eduardo Salas
Christopher F. Chabris	Ken Gilhooly	Raoul Oudejans	Dean Simonton
Neal J. Cohen	Fernand Gobet	Henning Plessner	Peter Verkoeijen
Nancy Cooke	Cleotilde Gonzalez	Damian Poulter	Gary L. Wells
Nelson Cowan	Dominik Guess	Robert Proctor	Dan Woltz
Keith Davids	D. Zachary Hambrick	Thomas Reilly	*Student Reviewer*
Reuven Dukas	Verlin Hinsz	Ian Renshaw	Andy Dattel