TITLE
Capturing and testing perceptual-cognitive expertise: A comparison of stationary and movement response methods

AUTHOR
Roca, André; Williams, A. Mark; Ford, Paul R.

JOURNAL
Behavior Research Methods

DATE DEPOSITED
10 August 2016

This version available at
https://research.stmarys.ac.uk/id/eprint/1160/

Capturing and Testing Perceptual-Cognitive Expertise: A Comparison of Stationary and Movement Response Methods

André Roca[1,2]*, Paul R. Ford[1], and A. Mark Williams[3]

[1]Research Institute for Sport and Exercise Sciences,

Liverpool John Moores University, UK

[2]Institute of Cognitive and Team/Racket Sport Research

German Sport University Cologne, Germany

[3]Centre for Sports Medicine and Human Performance,

Brunel University London, UK

*To whom correspondence should be addressed:

Institute of Cognitive and Team/Racket Sport Research

German Sport University Cologne

Am Sportpark Müngersdorf 6

50933 Cologne, Germany

Email: a.roca@dshs-koeln.de

Abstract

Numerous methods have been used to study expertise and performance. In the present article, we compare the cognitive thought processes of skilled soccer players when responding to film-based simulations of defensive situations involving two different experimental paradigms. Participants either remained stationary in a seated position (*n* = 10) or were allowed to move (*n* = 10) in response to life-size film sequences of 11 versus 11 open-play soccer situations viewed from a player's perspective. Response accuracy and retrospective verbal reports of thinking were collected across the two task conditions. In the movement-based response group, participants generated a greater number of verbal report statements, including a higher proportion of evaluation, prediction, and action planning statements than participants in the stationary group. Findings suggest that the processing strategies employed during performance differ depending on the nature of the response required by participants. Implications for behavioral methods and experimental design are discussed.

Keywords: *expert performance; representative task design; simulation fidelity; cognitive processes.*

Introduction

Expert performance has been examined across numerous domains, including aviation (e.g., Palmisano & Gillam, 2005), military combat (e.g., Williams, Ericsson, Ward, & Eccles, 2008), medicine (Kushniruk & Patel, 1998), and sport (Williams, Ford, Eccles, & Ward, 2011). Many of these researchers have used simulation in all its various guises, including virtual, computer, and film-based approaches, to recreate the performance environment under controlled and reproducible conditions in the laboratory (for a review, see Ward, Williams, & Hancock, 2006). In sport, the design of these representative tasks is a significant challenge for scientists attempting to faithfully capture and reproduce the performance environment, particularly the perceptual and cognitive demands of the task. The challenge has been to try and reproduce the highly dynamic and rapidly changing nature of the competition setting in a controlled, repeatable manner for experimentation.

Scientists have made attempts to use representative tasks with high fidelity, which is the degree to which a model or simulation reproduces the state and behavior of a ‘real-world’ feature or condition (Hays & Singer, 1989). However, the overriding tendency has been to design simplistic and sometimes manufactured or contrived tasks with a stronger emphasize on internal rather than external validity (Dhami, Hertwig, & Hoffrage, 2004). A concern is that such designs may introduce potential floor or ceiling effects on performance, denying experts access to information they would normally use and/or causing them to employ different processes to solve a particular task (Abernethy, Thomas, & Thomas, 1993).

Another important concern is the degree to which stimuli and their associated responses are related to each other or not, which is known as stimulus-response compatibility. Scientists exploring stimulus-response compatibility effects have shown that responses are faster and more accurate when a spatial stimulus array matches a spatial response array (compatible mapping) compared to when it does not match (incompatible mapping) (see Fitts & Deininger, 1954). Furthermore, when this compatibility is low the simple relationship between the potential amount of information to be processed and the capability to process it effectively can be affected (Proctor & Reeve, 1990). Thus, the stimulus-response compatibility effect appears to take place when there is a physical or conceptual similarity between the stimulus and response sets (Kornblum, Hasbroucq, & Osman, 1990).

In recent years, there have been concerns raised and much debate voiced about the representativeness and fidelity of research on expert performance (e.g., Dhami et al., 2004; Dicks, Davids, & Button, 2009; Ericsson & Williams, 2007). Researchers have become aware of the need to develop more representative experimental tasks for testing and training the processes and component skills underpinning expert performance. For example, in sport the fidelity of the experimental paradigm has been shown to influence the perceptual behaviors employed (Dicks, Button, & Davids, 2010). Dicks and colleagues (2010) showed that different and more pertinent visual search patterns were employed by experienced soccer goalkeepers under more representative task constraints compared with the less representative conditions. Therefore, there is a need to design and employ task-paradigms that provide realistic but reproducible domain-specific situations so that performance can be objectively evaluated over repeated tests.

In the current paper, we examine whether the fidelity of the response mode influences the underlying processing strategies governing anticipation and decision-making performance during a representative film-based simulation of 11 versus 11 open-play soccer situations. We use a novel approach to this issue by collecting retrospective verbal reports of thinking from skilled participants under two different response conditions that were either stationary or movement-based, in conjunction with standard measures of response accuracy. It was expected that the difference in response fidelity between the stationary and movement paradigms would lead to differences in the verbal reports articulated by participants across the two conditions (e.g., Proctor & Reeve, 1990). A greater proportion of higher-order verbal report statements of cognitive processes were expected, such as predictions and action planning, for the movement compared with the stationary response group due to the higher fidelity of the representative task design and increased stimulus-response compatibility. Moreover, we hypothesized that participants in the movement group would demonstrate more accurate performance in comparison with their counterparts in the stationary group (e.g., see Fitts & Deininger, 1954).

Material and Methods

Participants

In total, 20 male semi-professional soccer players participated in the experiment. Participants were randomly allocated into two different experimental groups: movement-based (*n* = 10) and stationary (*n* = 10). The movement group took part in a more realistic representative task that included an action component, while in the stationary group a less realistic verbal response task was employed with participants remaining in a seated position. Participants in the movement group (*M* age = 21.5 years, *SD* = 2.0) had played soccer regularly since the mean age of 5.6 years (*SD* = 1.2), during which they had trained/played for a mean of 9.2 hr (*SD* = 1.7) per week and participated in an average total of 615 (*SD* = 131) competitive matches. The stationary group (*M* age = 21.1 years, *SD* = 2.0) had regularly participated in soccer since the mean age of 5.8 years (*SD* = 1.6), during which they had trained/played for a mean of 9.0 hr (*SD* = 1.8) per week, including participation in an average total of 632 (*SD* = 145) competitive matches. Informed consent was provided prior to participation and ethical approval was gained through the lead institution's Ethics Board.

Stimuli and apparatus

Participants were presented with life-size video sequences involving dynamic, 11 versus 11 soccer situations filmed and viewed from the perspective of a central defender and with the opposition team in possession of the ball (for further details on the production of the video-based test stimuli, see Roca, Ford, McRobert, & Williams, 2011). The video stimuli consisted of 20 test and four practice trials. All video clips were approximately 5 s in duration, with each one being occluded at a key moment in the action, such as when the opposition player in possession of the ball was about to make a pass, shoot at goal, or maintain possession of the ball by dribbling forward.

The test film stimuli were back projected onto large projection screen (Draper Cinefold, Spiceland, IN, USA; height, 3 m, width, 4 m) that was placed at a distance of 3 m directly in front of the participant. Participants in the movement group were free to move and interact with the action sequence as they would normally do when playing in a real soccer match, which includes moving forward, backwards, and sideways (see Figure 1). The movement of participants were monitored using a digital video camera (Canon LEGRIA FS200, Tokyo, Japan) positioned 3m behind the participant and linked to a TV monitor screen (Philips 15PF5120, Eindhoven, Netherlands) placed on the experimenter's desk. In contrast, in the stationary group, participants were seated during the experiment at the same start position as in the movement group.

Verbal responses and reports of thinking were collected using a lapel wireless microphone system (Sennheiser EW-100G2, Wedemark, Germany), including a telemetry radio transmitter fixed to the participant and a telemetry radio receiver connected to the digital video camera.

Procedure

Prior to the experimental task, participants were given instruction and training on how to think aloud and provide retrospective verbal reports. The instructions comprised Ericsson and Kirk's (2001) adaption of Ericsson and Simon's (1993, pp. 375-379) original protocol combined with a series of domain-specific warm-up tasks. Training continued until participants were comfortable with the procedure and feedback was given to ensure that participants' verbal reports were consistent with the instructions. The verbal report training protocol lasted approximately 30 min. Once training had been completed, participants were presented with a total of four warm-up trials to ensure familiarization with the experimental setting and the protocol procedure. Retrospective verbal reports were collected directly after every trial. After providing verbal reports, the participants were required to verbally confirm 'What the player in possession was going to do?' and 'What decision the participant themselves made or were about to make at the moment of video occlusion?' Participants completed 20 test trials in a quiet room and each individual completed the training and test session in about 60 min.

Two outcome measures of performance were obtained. Anticipation accuracy was defined as whether or not the participant correctly selected the next action of the player in possession of the ball at the moment of video occlusion such as a pass to a teammate, a shot at goal, or continued dribbling the ball. A panel of three Union of European Football Associations (UEFA) qualified soccer coaches independently selected the most appropriate decision for a participant to execute in response to the on-screen situation at the time of video occlusion on each trial. The inter-observer agreement between coach selections was 91.7%. Decision-making accuracy was defined as whether or not the participant decided on the action selected by the coaches as most appropriate for that trial. The correspondence between the movement group's action selection (as determined through verbal confirmation by participants of their decision at the end on each trial) and action execution (as determined through video observation of participants on each trial) was 100%. Anticipation and decision-making accuracy were calculated as the mean number of trials (in %) in which the participant selected the correct response.

The verbal report data were analyzed using the three most discriminating trials between groups, which were chosen based on the mean scores from the anticipation and decision-making measures (cf., McRobert, Williams, Ward, & Eccles, 2009). Participants' retrospective verbal reports were transcribed verbatim and segmented using natural speech and other syntactical markers. Verbal reports were classified according to a structure adapted from Ericsson and Simon (1993) and further developed by Ward, Williams, and Ericsson (2003). Four major types of cognitive thought statement categories were coded: (a) *monitoring statements* were those recalling current actions or descriptions of current events; (b) *evaluations* were statements making some form of comparison, assessment, or appraisal of events that are situation, task, or context relevant; (c) *predictions* referred to statements anticipating or highlighting future or potential future events; and (d) *planning statements* were those referring to a decision(s) on a course of action in order to anticipate an outcome or potential outcome. The reliability of the data was established using the intra- (94.5%) and inter-observer (93.3%) agreement formulas. These figures were created from a re-analysis of 20.0% of the data using procedures recommended by Thomas, Nelson, and Silverman (2005).

Results and Discussion

Independent *t*-tests showed no significant difference in anticipation response accuracy between the movement-based (M = 60.0%, SD = 12.9) and stationary groups

(*M* = 59.5%, *SD* = 7.3), $t(14.02) = 0.11$, $p = .92$, $d = 0.05$. However, there was a trend towards significance for the decision-making response accuracy scores, $t(18) = 1.79$, $p = .090$, $d = 0.80$. The mean percentage of correct decision-making responses for the movement group (*M* = 79.5%, *SD* = 8.0) was slightly higher than that for the stationary group (*M* = 73.0%, *SD* = 8.2). Differences in fidelity between the two groups although did not have a significant impact on the accuracy of the judgments made.

A 2 Group (stationary, movement) x 4 Statement Type (monitoring, evaluation, prediction, planning) ANOVA revealed a main effect for group, $F(1, 18) = 280.33$, $p < .001$, $\eta_p^2 = .94$. The movement group (*M* = 7.37 statements, *SD* = 2.05) generated significantly more verbal statements of cognitive processes when compared with the stationary group (*M* = 4.37 statements, *SD* = 0.85). There was also a significant effect for type of verbal statement, $F(3, 54) = 61.40$, $p < .001$, $\eta_p^2 = .77$. Bonferroni-corrected pairwise comparisons demonstrated that participants verbalized significantly more monitoring statements (*M* = 3.27 statements, *SD* = 1.12) compared with all other statement types. A higher number of predictive statements (*M* = 1.15 statements, *SD* = 0.44) were verbalized compared with evaluation statements (*M* = 0.60 statements, *SD* = 0.71). No differences were found between planning statements (*M* = 0.85 statements, *SD* = 0.81) and evaluation or predictive statements. These data are presented in Table 1.

The Group x Type of Verbal Statement interaction was not significant, $F(3, 54) = 0.58$, $p = .63$, $\eta_p^2 = .03$. However, because the movement group made more statements compared to the stationary response group the frequency scores for each category were subsequently normalized into percentage data. Participants in the stationary group made a greater proportion of monitoring statements compared to any other type of statement (*M* = 67.4%, *SD* = 13.0 vs. *M* = 32.6%, *SD* = 13.0). In contrast, participants in the movement group made a lower proportion of monitoring statements compared to any other statement type (*M* = 48.6%, *SD* = 10.1 vs. *M* = 51.4%, *SD* = 10.1), indicating they engaged in a greater amount of higher-order cognitive thought processing involving evaluations, predictions, and planning.

Our findings support the hypothesis that differences in the fidelity of the two response paradigms (Hays & Singer, 1989) would lead to differences in the verbal reports articulated by participants. Moreover, it supports previous research that shows that the capability to process information effectively is reduced when there is a lower 'natural' connection between the stimulus and the required/associated response (e.g., Proctor & Reeve, 1990). The cognitive thought processes observed for the movement-based group may be a result of the improved compatibility between stimulus characteristics and response selection/execution under the more realistic settings. That is, the need to move in response to the continuous action presented on the life-size screen appears more compatible with the skilled player's customary response in the game situation. A further interpretation could be that participants in the movement-based response group may have invested more mental effort in the task and felt more engaged due to the increased fidelity (e.g., Bianchi-Berthouze, 2013). In future, a rating scale for mental effort (e.g., RMSE: Zijlstra, 1993) could be used to measure the participants' perceived engagement and investment on the different task conditions.

Findings highlight the importance of designing representative tasks that offer participants more realistic context for continuous decision making, perception, and action as per the environmental characteristics of the actual performance domain. Such tasks are preferable to those that isolate each/any of these elements of performance (Dicks et al., 2010; Ericsson & Williams, 2007; Ward et al., 2006). The suggestion is that greater attention needs to be displayed to the fidelity and ecological representativeness of task designs so that inferences and conclusions can be made about the specific and often complex processes that underpin and mediate expert performance.

In conclusion, we examined whether the cognitive strategies employed during filmed simulation differed depending on whether participants remained stationary or moved/interacted with it as they normally would. Participants in the movement group verbalized a larger number of thought statements with a higher proportion related to the assessment and prediction of future options and the planning and selection of an appropriate action response when compared with the stationary group. The higher fidelity and greater stimulus-response compatibility evident in the movement group led to different thought processes being engaged when compared to the stationary group, albeit these changes did not have a marked impact on the accuracy of the judgments made. Our findings highlight the need to design experimental paradigms that (more closely) recreate the constraints that exist in the actual performance setting in order to better identify the mechanisms and processes mediating superior performance.

References

Abernethy, B., Thomas, K. T., & Thomas, J. T. (1993). Strategies for improving understanding of motor expertise (or mistakes we have made and things we have learned!). In J. L. Starkes & F. Allard (Eds.), *Cognitive issues in motor expertise* (pp. 317-356). Amsterdam, Netherlands: Elsevier.

Bianchi-Berthouze, N. (2013). Understanding the role of body movement in player engagement. *Human-computer interaction, 28*, 40-75.

Dhami, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin, 130*, 959-988.

Dicks, M., Button, C., & Davids, K. (2010). Examination of gaze behaviors under in situ and video simulation task constraints reveals differences in information pickup for perception and action. *Attention, Perception, & Psychophysics, 72,* 706-720.

Dicks, M., Davids, K., & Button, C. (2009). *Representative task designs for the study of perception and action in sport. International Journal of Sport Psychology, 40,* 506-524.

Ericsson, K. A., & Kirk, E. (2001). *Instructions for giving retrospective verbal reports*. Tallahassee, FL: Florida State University.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: Bradford books/MIT Press.

Ericsson, K. A., & Williams, A. M. (2007). Capturing naturally occurring superior performance in the laboratory: Translational research on expert performance. *Journal of Experimental Psychology: Applied, 13,* 115-123.

Fitts, P. M., Deininger, R. L. (1954). S–R compatibility: correspondence among paired elements within stimulus and response codes. *Journal of Experimental Psychology*, *48*, 483-492.

Hays, R. T., & Singer, M. J. (1989). *Simulation fidelity in training system design: Bridging the gap between reality and training*. New York, NY: Springer.

Kornblum, S., Hasbroucq, T., Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility - A model and taxonomy. *Psychological Review, 97,* 253-270.

Kushniruk A. W., & Patel V. L. (1998). Cognitive evaluation of decision making processes and assessment of information technology in medicine. *International Journal of Medical Informatics, 51,* 83-90.

McRobert, A. P., Williams, A. M., Ward, P., & Eccles, D. W. (2009). Tracing the process of expertise in a simulated anticipation task. *Ergonomics, 52,* 474-483.

Palmisano, S., & Gillam, B. J. (2005). Visual perception for touchdown during simulated landing. *Journal of Experimental Psychology: Applied, 11*, 19-32.

Proctor, R. W., & Reeve, T. G. (Eds.) (1990). *Stimulus-response compatibility: An integrated perspective*. Amsterdam, Netherlands: Elsevier.

Roca, A., Ford, P. R., McRobert, A. P., & Williams, A. M. (2011). Identifying the processes underpinning anticipation and decision-making in a dynamic time-constrained task. *Cognitive Processing*, *12*, 301-310.

Thomas, J. R., Nelson, J. K., & Silverman, S. J. (2005). *Research methods in physical activity* (5th ed.). Champaign, IL: Human Kinetics.

Ward, P., Williams, A. M., & Ericsson, K. A. (2003). Underlying mechanisms of perceptual-cognitive expertise in soccer. *Journal of Sport & Exercise Psychology, 25*, S136.

Ward, P., Williams, A. M., & Hancock, P. A. (2006). Simulation for performance and training. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 243-262). New York, NY: Cambridge University Press.

Williams, A. M., Ericsson, K. A., Ward, P., & Eccles, D. W. (2008). Research on expertise in sport: Implications for the military. *Military Psychology, 20*, S123–S145.

Williams, A. M., Ford, P. R., Eccles, D. W., & Ward, P. (2011). Perceptual-cognitive expertise in sport and its acquisition: Implications for applied cognitive psychology. *Applied Cognitive Psychology*, *25*, 432-442.

Zijlstra, F. R. H. (1993). *Efficiency in work behavior: A design approach for modern tools*. Delft, Netherlands: Delft University Press.

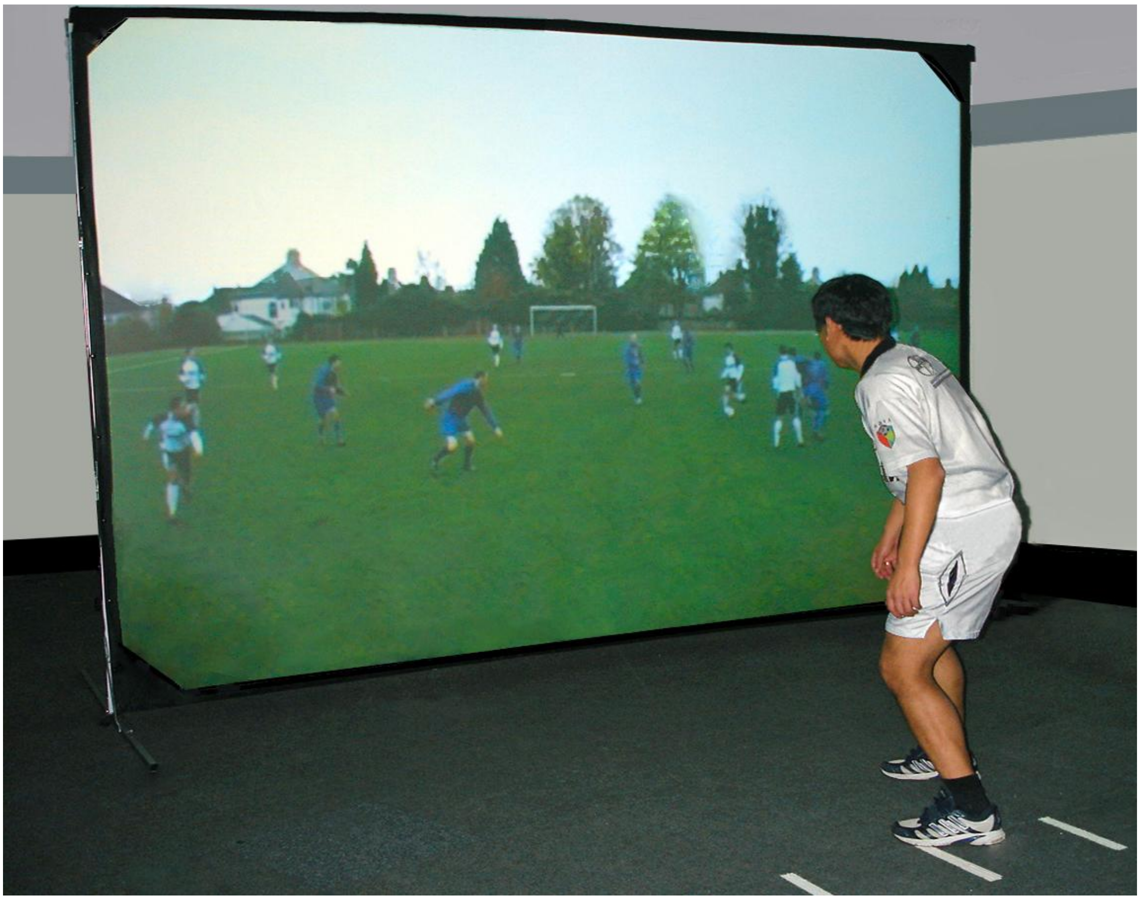

Figure 1. The experimental layout employed in the movement-based response paradigm.

Table 1

Mean (SD) Number of Type of Verbal Statements Generated per Trial Between Stationary and Movement Response Groups

	Verbal statement category			
Group	Monitoring	Evaluation	Prediction	Planning
Stationary response	2.93 (0.83)	0.27 (0.47)	0.87 (0.36)	0.30 (0.48)
Movement response	3.60 (1.30)	0.93 (0.77)	1.43 (0.32)	1.40 (0.70)