Towards a theoretically-driven model of correspondence between behaviours in one context to another: Implications for studying sport performance

DUARTE ARAÚJO* and KEITH DAVIDS**/***

(*) Spertlab, CIPER, Faculdade de Motricidade Humana, Universidade de Lisboa, Portugal
(**) Centre for Sports Engineering Research, Sheffield Hallam University, UK
(***) FiDiPro Programme, Faculty of Sport and Health Sciences, University of Jyväskylä, Finland

Abstract

Introduction

Among her many accomplishments in a distinguished career, over two decades ago Janet Starkes proposed an important change with significant consequences for empirical research in sport psychology. Her proposal was that the use of technology “may help us to move away from purely laboratory simulations to test situations with more realistic perceptual information” (Starkes, Edwards, Dissanayake & Dunn, 1995, p.166). This proposal is completely aligned with key ideas in ecological dynamics, espousing the relevance of studying sport performance at the individual-environment scale of analysis. The use of terms like “in situ” (e.g., Farrow & Abernethy, 2003), “on-court” tests (Starkes & Lindley, 1994), “field” tests (Starkes et al., 1995) and the like, capture how sport psychology research, influenced in part by
Starkes’ contributions, moved away from the confines of conventional psychology, with its traditional indifference to the nature of the constraints in specific performance ecologies outside the laboratory. This paradigm shift (see Travassos et al. 2013, for a meta-analysis) echoes the words of Abernethy (1987), who stated that “In view of the need to preserve ecological validity within the testing environment (Neisser, 1976) a more appropriate approach to examining anticipatory ability in fast ball sports than that provided by the use of basic laboratory tasks would appear to be provided by experimentation using sport-specific stimuli” (p.7).

Starkes’ important contribution to sport psychology highlighted the functional relevance of in situ methodologies for data collection, effectively proposing the use of sport performance contexts as a type of richer (field) laboratory setting (Starkes et al., 1995). However, with a few exceptions (e.g., Pinder et al., 2011a), sport psychologists have not sought en masse to develop a principled way to “preserve ecological validity” (Abernethy, 1987, p.7) in sport experiments. This may be due to a misconception of the term “ecological validity”, which originated with Neisser (1976) himself. Neisser (1976) acknowledged that Egon Brunswik had coined the concept “ecological validity”, but at the same time he mentioned that Brunwik’s “usage was slightly different from the one that is popular today.” (footnote on page 48, for the text written on page 33), without clarifying the definition of the term. Twenty years earlier, Brunswik (1956) defined ecological validity in test measurement terms. Neisser mistook the term ‘ecological validity’ for the concept of ‘representative design’, but gave some much-needed impetus to concerns on empirical methods, as expressed by Starkes and colleagues.

Thus, it is with no surprise that Starkes’ advocacy of in situ designs also raised many pertinent questions related to principled criteria to classify an experimental task design as in situ. After all, to achieve this level of task representativeness in investigating sport performance, it might be (il)logically construed that experimenters need to only study motor behaviours observed within competitive performance environments. This is clearly not possible, nor desirable, due to the presence of irrelevant idiosyncrasies of specific competitive events which might contaminate data. What does this conceptualisation of in situ designs imply for laboratory experiments on sport performance? Should they be avoided because of their potential reductionism in not studying only behaviours in specific competitive events? As we discuss later, the notion that the theoretical assumptions of ecological psychology preclude the use of controlled experimental conditions is a misguided view. Therefore, a clarification of a key tenet in ecological psychology is re-presented: That the relationship between information and action underpins per-
formance in studies of human behaviour. In this position paper we argue that following some principles, this relationship can be maintained in the experimental laboratory, in the competitive environment and even in virtual environments, and therefore, behavioural correspondence between contexts can be evaluated.

It is important to highlight that terms like *in situ* are normally included in research articles to mean ‘not in the laboratory’, the implication being that ‘ecologically valid’ is ‘good for sport sciences’. In this position paper we argue that actually terms like *in situ*, while useful, are somewhat operational in nature rather than theoretical. Indeed, typically, the implementation of an *in situ* experimental task design is predicated on mainly dichotomous methodological arguments about preservation of internal and external validity. Here we argue that what is needed now is a focus on theoretical principles which guide the (re)design of experimental task constraints that allow functional behaviours to emerge from participants. These functional behaviours can emerge under carefully designed task constraints in the laboratory, in simulation environments or in performance contexts, and, indeed, anywhere in between, as long as key theoretical principles guide empirical (representative) design.

In order to deconstruct the misconception that laboratory experiments should be avoided because they are not “ecologically valid”, here we comment on this specific contribution of Janet Starkes to sport psychology methodology, by clarifying the concepts of representative design and ecological validity, and presenting a theoretical framework for the criteria to make a correspondence between behaviour in one context to behaviour in another context.

**In situ Experiments Highlight The Need For Representative Design**

Nearly a decade ago we published an experiment on decision making in the sport of sailing, where the task for participants was to perform in a computerised, simulated regatta (Araújo, Davids & Serpa, 2005). The conceptual foundations of the study were based on previous findings in cognitive psychology that have demonstrated the external validity of performing on computer simulations (e.g., Sanders, 1991, Brehmer, 1996, Brehmer & Dorner, 1993), where external validity is the degree to which the results can be generalized to other samples and situations. In our case, external validity was obtained with a high correlation between the level of expertise of the sailors and their performance on the sailing simulator. However, our design of this
2005 study led to a series of rich discussions, not least with anonymous reviewers of our journal submissions, that helped us gain a clearer understanding of the theoretical rationale behind the construction of an experimental task. This theoretical interpretation led us to elaborate on criteria for the correspondence of the design of our experimental task constraints with behavioural phenomena towards which we sought to generalise. This is a key feature of experimentation in psychology that had been originally proposed many years ago by Egon Brunswik (1956), with his methodology of “representative experimental design” (see Dhami et al., 2004 for a complete overview). In proposing this idea, Brunswik (1956) defended the view that cues or perceptual variables should be sampled from an organism’s environment so as to be representative of the environmental stimulation to which it had adapted, and which formed the focus of an experimenter’s generalization.

Essentially, the phrase representative experimental design refers to the arrangement of conditions of an experiment so that they represent the behavioural setting to which the results are intended to apply. Hammond and Stewart (2001) noted that Brunswik used the term represent here in the same sense in which a sample of participants in an experiment might be said to represent individuals in a particular population that is a focus of study in an experimental programme. Hammond and Bateman (2009b) provided a good illustration of this point. They commented on a paper by Williams and colleagues (2002), which presented data on anticipation of an opponent’s actions in tennis. Although Williams et al. were careful in sampling two groups of participants: highly skilled and less skilled, only skilled models were used in images to simulate actions of opponents in these experiments. As Hammond and Bateman (2009b) argued, the stimuli for less skilled players were not represented. Indeed, no criteria were provided to help us understand whether the results of these experiments would generalize to explanations of behaviour in a performance ecology where less skilled players compete with other less skilled players. In effect, Brunswik (1956) was arguing that the (statistical) logic of induction in experiments should hold for sampling from behavioural environments, as well as from specific populations of participants.

Hammond and Stewart (2001) emphasized that Brunswik did not mean that all experiments must represent the conditions of the ‘real’ world by means of statistical sampling (see Hammond, 1998, Stoffregen, 1993, for a discussion around the vagueness of terminology like “real,” “typical,” or “natural” world). Rather, Brunswik argued that the logic of induction should be applied in both directions (from environment to experiments and back),
if a generalization of findings is to be justified. Indeed, he often used the expression double standard to indicate that psychologists used the logic of induction in one direction only (over participants), but typically claimed generalization over environments without justification.

According to Brunswik, in experiments on behaviour, individuals need to be required to cope with the multiple, noisy, messy situations, which occur in specific performance environments. Only by representing those irregular conditions to a behaving individual can psychologists discover how he/she achieves a patterned relation with a specific performance environment, despite the apparent uncertainty engendered.

The critical point Brunswik raised was that experimenters should avoid the tendency to oversample highly improbable perceptual variables in an intended behavioural setting, for example avoiding use of artificially created perceptual variables (e.g., use of LEDs coloured red or green to signal which direction athletes should run towards or aim a ball towards). Such informational constraints do not exist in sport performance environments (e.g., team sports). Brunswik argued that the description of a task in testing hypotheses should be provided by estimating the ecological validity of cues (this term is correctly defined as the degree of correlation between a proximal cue and the distal variable to which it is related, see Brunswik, 1956, pp. 48-52; see also Araújo & Kirlik, 2008, Hammond & Bateman, 2009a and b, for a detailed explanation in sport about proximal and distal cues).

For example, in a perceptual task, ecological validity refers to the precisely measured correlation between the perceived size of an object (proximal cue) and distance of the object from the observer (the distal cue, or the criterion) over a series of performance situations, the intercorrelation among these cues, and the overall uncertainty in the task. This is a formal way to quantify the correspondence between experimental task constraints and a represented behavioural setting. For instance, cue number, values, and distributions might range from 0 to infinity, and the ecological validities of cues and their intercorrelations range from −1 to +1 (for a detailed discussion on Brunwikian correlational methods, see Cooksey, 1996). Specific samples of situations in behavioural settings would lie within these boundaries.

A researcher who uses statistical situation sampling can sample various combinations of environmental properties (e.g., various numbers of cues, ecological validities, and intercue correlations). To exemplify, Jacobs, Michaels and Runeson (2000), investigated how participants perceived the relative mass of colliding balls, by manipulating the perceptual variables available to participants (i.e., the distribution of four cues: scatter angle, exit speed, pre-collision speed and mass ratio invariant), with different values for each vari-
able. Different participants used different variables for judgments of mass. They also observed during training that many perceivers changed which variables (or combination of variables) they used, and that after a minimal amount of training, at least some participants seemed to detect the cues with the highest ecological validity (with +1 defining a highly reliable relationship between proximal and distal cues, and -1 capturing the inverse; for other examples see Cooksey, 1996, Hammond & Stewart, 2001). This was the method Brunswik used to develop task constraints to capture and reproduce environments in a representative manner in controlled experiments on human behaviour.

After Brunswik, other authors proposed theoretical frameworks for understanding representativeness (Hoffman & Deffenbacher, 1993), criteria to generalize transfer (Barnett & Ceci, 2002), or empirical demonstrations of transfer (particularly in sport see Abernethy, 1991, Farrow & Abernethy, 2003; Williams et al., 2002; see also Risko, Laidlaw, Freeth, Foulsham & Kingstone, 2012, for an empirical approach to the determination of the equivalence of different social stimuli). However, these considerations have not generated operational principles for research design. Indeed, the main problem has been maintained: The need to justify how the specific conditions under which the results of an experiment are obtained can represent the performance conditions to which the results will be applied. Hammond and Bateman (2009a) argued that researchers need to make explicit the particular features of the performance situations towards which the results of an experiment are intended to apply. Making those features explicit implies analysing one’s theory, and denoting those features of the environment that are theory-relevant. Building on Brunswik’s (1956) initial attempt, we propose an empirical approach that broadens the concept of representative design to that of behavioural correspondence between performance contexts.

**Behavioural correspondence between contexts**

The measurement of “Brunswik’s” representativeness can be very difficult in many situations, such as those of direct competition with an opponent or opposing team in sport (e.g., Cooksey, 1996, Araújo et al., 2007). For example, how is it possible to know in advance the composition of an adversary team, what the “mission” of each adversary player is for the game, and how the players are interrelated for the specific match? Sport analysts may speculate, but they cannot calculate the precise ecological validity of cues, their interrelation, and the overall uncertainty of a future competitive game, without knowing which cues will be perceived by performers.
Another point that demands an approach that goes beyond Brunswik’s “representative design” is that by randomly sampling cues in an environment, the researcher may be excluding those unique cues that are available for attuned, expert participants. This means that, more than sampling cues in an environment, it may be more relevant to select affordances (i.e., possibilities for action). For Brunswik, how the individual acts was not very relevant. Brunswik, as well as his influential co-author Edward C. Tolman (e.g., Tolman & Brunswik, 1935) clearly rejected the so-called molecular approaches and the particular behaviours of organisms were not a key issue. What mattered for them were the consequences of the behaviour (the molar level), i.e., what the organisms were getting at, not how they were getting there (see also Goldstein, 2004, p.40). For sport scientists, as well as for Gibsonian ecological psychologists, the problem of “how to act” in a given situation is a major concern, given that the structure of action is an expression of goal-directness (e.g., Turvey, 2013).

The representative task as a “modified environment” implies a specific adaptation of the athlete, since both contexts (the representative task and the performance context) may have different affordances (Araújo et al., 2007, Pinder et al., 2011a). This implies a theoretical change, from understanding an environment that interacts with a performer (i.e., two separate systems that interact) to a more holistic view, which understands performer and environment as a single coherent system to be studied. For Gibson (1979, see also Richardson et al., 2008) this is the ecological level of analysis where organisms and behaviour are defined; an ecology cannot be defined without studying the organisms it comprises, and vice-versa.

Finally, there are contexts which are not representative (from a Brunswikian point of view), or even which may never be observed in a competition (the context towards which one wants to generalize), but which can contribute to understanding and testing the performance of athletes in competition. For example, in team sports like futsal or basketball, coaches in some specific circumstances prepare their five-player teams with overload training against teams with 6 and 7 players. In competition this will never occur, but coaches use this method to train their players to work harder to create space away from opposing players, for a more rapid and accurate perception of passing gaps, for dribbling past immediate and covering defenders, and for shooting under severe defensive pressure. The overloading of opposing players is undertaken in order for trainees to become better perceptually attuned to affordances that are relevant in competitive performance environments. We may think about similar non-representative contexts built on virtual environments, where researchers (and coaches) may...
manipulate circumstances relevant to understand performance in a competition, but that could not be manipulated out of the virtual context (e.g., Correia et al., 2012). In these particular contexts researchers can measure perception, action, and cognition in the same tradition of Starkes and her colleagues.

These limitations of Brunswik’s concept of representative design led us to advance the concept of behavioural correspondence between contexts (see Araújo et al., 2007; Davids et al., 2012). In particular, we seek to propose principles to demonstrate the correspondence between behaviour in an experimental setting and behaviour in a context towards which we want to generalize a behaviour measured in an experimental setting. This experimental setting can be a traditional laboratory, a virtual laboratory or an in situ performance context, envisaged by Starkes.

The interrelated criteria to test behavioural correspondence between contexts are:

1) Selection of functional (i.e. relevant) affordances. Even though affordances can be empirically tested (see Fajen, Riley & Turvey, 2009; Richardson et al., 2008, for reviews), their selection should be theoretically driven, even if the researcher is not adopting an ecological psychology standpoint. The idea is that the selected action possibilities (affordances) could be perceived and acted upon by a performer. This point was well illustrated by Abernethy and colleagues (2001):

“While the findings of Starkes et al (1995) are reassuring with respect to the generalizability of conclusions that can be drawn from laboratory temporal occlusion studies to anticipation in a natural task setting, and the use of the rich natural perceptual display represents an important methodological advance from film simulation, perception in the task used by Starkes et al was still passive and completely decoupled from the normal action it informs and is informed by. This is an important limitation given the essential transactional relationship between perception and action… , and the convincing demonstrations that the full capabilities of perception can only be realized if perception is examined in situations where its natural couplings to action are retained…” (p. 242)

Functionality of affordances is also related to the actions they invite from performers, as well as the performance success they support. The next two criteria test the adequacy of this selection.

2) Action fidelity. Stoffregen et al. (2003) discussed the concept of “action fidelity” for computer simulations and virtual environments. For them the key aspects of this concept are that: i) perception is defined with respect to behaviour, and ii), action fidelity does not mandate a concentration on “stimulus fidelity” (in contrast with other attempts like for example, Risko,
et al., 2012), since the environment is defined in behavioural terms (i.e., affordances). Therefore, action fidelity concerns the degree to which actions performed in the experimental setting are related to the actions performed during competition. It can be measured in terms of task-related measures, such as variance in performance across trials, trials to criterion, or spatiotemporal patterns of interaction with others, objects, surfaces, gaps, terrain, etc (Araújo et al., 2007; Travassos et al., 2012, see Figure 1). Importantly, action fidelity is aligned with performance achievement (the next criterion), since creative action (i.e., providing a lower correlation with an action solution previously observed in the competitive environment) is acceptable if it emerges under the constraints of the competitive performance environment (i.e., under the rules of the game) and if it achieves a goal, such as to score, prevent scoring, pass the ball, intercept, dribble, cover and the like.

3) Performance achievement. Achievement is the degree of success obtained when performing a task for a specific goal. It can be seen as a concept with similitudes with that of ecological rationality, which was developed by Gigerenzer and refers to the degree of effectiveness of a specific action (see Hammond, 2007, for a discussion). For Gigerenzer, behaviour is successful if it is adapted to the structure of the information in the environment in which it is realized (Gigerenzer, Todd & ABC research group, 1999). That is, the ecology is the reference point for the observation of a behaviour. This view is contrasted with traditional normative references, which define a priori what a classical movement technique is or the “best” solution according to a social definition, even when it is not effective. This can be exemplified in methods for training a classical sport technique, such as shooting at the basket, even if a performer is missing the basket or practising a classical driving technique on a golf range under the same conditions (the aim to repeat a spe-

![Fig. 1. - Two examples of speed variations for a pre-determined pass (left panel) and an emergent pass (right panel) (From Travassos et al., 2012).](image)
cific movement pattern, not to reduce the distance of the ball to a certain target). Achievement can be assessed with output task measurements (ball entering the basket or distance of the drive/closeness to the green in golf).

Recently, our group has tested the idea of behavioural correspondence between contexts by including affordances that are theory-relevant, spatiotemporal patterns of interaction to measure action fidelity and the criteria to define success (performance achievement) (see Pinder et al., 2011b, Travassos et al, 2012). Pinder and colleagues’ (2011b) results on a cricket batting task showed that each distinct set of task constraints led to the emergence of a specific, functional coordination solution. Removal of advanced information sources from a bowler’s actions, when the batters faced a ball projection machine, caused significant delays in movement initiation, and resulted in reduced peak bat swing velocities and a reduction in the quality of bat–ball contact, compared to batting against a “live” bowler. When responding to a two-dimensional video simulation, batters were able to use information from the bowlers’ action, enabling fidelity of initial behavioural responses consistent with the task of batting against a “live” bowler. However, without interceptive task requirements or actual ball flight information, significant variations in downswing initiation timing and peak bat velocities were demonstrated. Panchuk et al. (2013) observed similar variations in emergent behavioural responses when informational constraints on a one-handed catching task were changed. In another example, Travassos and colleagues (2012) in the team sport of Futsal, required players to perform a passing task in which uncertainty of passing direction for the player in possession of the ball was increased under four distinct conditions and compared with passing data observed during a competitive match. Higher levels of regularity were observed in predeterm ined passes (distance and direction of pass verified) compared to passes made under practice conditions which included more uncertainty for performers (i.e. distance, angle and direction of pass varied). Importantly, passing speed regularity and accuracy were more similar between practice tasks with higher levels of uncertainty and during competitive performance (Figure 1). The data convincingly demonstrated that only the design of the more uncertain passing tasks represented the speed, accuracy and success found in passing circumstances performed during a competitive match. These data show how the informational constraints of practice tasks can be designed to correspond to the informational constraints of a competitive performance environment in team sports.

In conclusion, Starkes’ ideas on the representativeness of in situ designs drew the attention of sport psychologists to the design of experimental conditions during performance in sport-related tasks. Concern over the rela-
The relationship between experiments and performance environments goes beyond the longstanding tension between internal and external validity, as Brunswik showed a long time ago. Brunswik developed important theoretical principles, which can be integrated into a broader ecological dynamics view enhancing our understanding about how behavior in one context may correspond to behavior in another context. The suggestion is that modeling the “behavioral correspondence between contexts” ensures that such generalizations can be achieved, based on adherence to the intertwined criteria of selection of affordances, action fidelity and performance achievement.

REFERENCES


