ORIGINAL ARTICLE
Topic: Event Predictive Cognition

Action Production and Event Perception as Routine Sequential Behaviours

Richard P. Cooper

Correspondence
Richard P. Cooper, Centre for Cognition, Computation and Modelling, Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London, WC1E 7HX, United Kingdom
Email: R.Cooper@bbk.ac.uk

Funding information
This work was not funded through any specific grant

It is argued that event perception and routine sequential action production share a range of characteristics (e.g., similar levels of automaticity, the involvement of sequentially and hierarchically organised schemata, and the coupled operation of predictive and monitoring processes). With this in mind, and in an effort to develop a mechanistic account of event perception, we consider how an existing model of routine sequential action production might be applied in the domain of event perception. We focus the discussion on the multiple roles of prediction in the two domains, and consider the implications of the application of the model of action production to event perception, and for sequential processing more generally.

KEYWORDS
Action control, event perception, monitoring, prediction, error correction, action schemata, event schemata, contention scheduling

1 | INTRODUCTION: ON THE RELATION BETWEEN EVENTS AND ACTIONS

It has been convincingly argued that in order to make sense of the ever changing world we structure perceptual experience in terms of events, drawing upon schematic or abstract knowledge of different types of events in order to facilitate the process (Zacks and Tversky, 2001). Within this conception, an *event* is a coherent series of temporally or causally connected interactions culminating in some agent-relevant regularity or goal, while an *event schema* is a representation of the abstract shared features of a set of similar events. Through perception, sequences of simple or atomic changes in the perceived world give rise to or activate an instance of an event schema, resulting in a mental representation of an event that ties the perceptual changes into a coherent whole. It is this coherent whole that allows the perceiver to make sense of the ongoing sequences of changes in the perceived world.

Event schemata are central to the process of event perception. Like schemata in other domains such as long term memory (Bartlett, 1932), motor control (Schmidt, 1975), and action control (Norman and Shallice, 1986), and as in the domain-general schema-theory of Arbib (1992) (see also Cooper, 2016), event schemata are abstractions over instances (in this case, instances of events) that capture functionally relevant regularities of those instances. Such regularities include the generic subcomponents or basic actions of the event (sometimes referred to as "scenes"), the sequential or temporal order of those subcomponents, and the culmination or goal of the event schema (Zacks and Tversky, 2001). Hierarchical relations also exist both over the components of events / event schemata (partonomic relations), and over types of event schemata (taxonomic relations). For example, a *football game event* will typically include subcomponents such as a *kick-off event* at the beginning of each half, and a *half-time event*, while the *football game event* is itself a specific kind of *team field sporting event*.

One key aspect of the sequential / temporal organisation of the components within event schemata is that it allows them to support prediction and event segmentation. Prediction is supported because once a scene has been identified as part of a known event schema, one may expect that subsequent scenes will correspond to subsequent components of the schema. Event segmentation is supported because at the end of an event, uncertainty or prediction error will be high (Kurby and Zacks, 2008).

A further property of event perception relevant to the current discussion is that it appears to be an automatic process, proceeding with little overt attention or conscious control (Zacks and Swallow, 2007). That is, as an observer we perceive a series of connected scenes as a coherent event. We cannot choose not to perceive the event.

In each of the above respects the domain of event perception parallels that of the production of routine sequential action. Many everyday activities involve the production of ordered sequences of basic actions, with those actions forming partonomies within temporally extended behaviour (e.g., preparing a cup of tea being part of preparing breakfast) and taxonomies with respect to more general types of action (e.g., preparing a cup of tea being a specific type of beverage preparation schema). Moreover routine sequential action has been argued to be goal directed (Cooper and Shallice, 2000) and automatic (Norman and Shallice, 1986). Finally, as argued below, the schematic nature of the representations that support routine sequential action production (*action schemata*) may also support predictive processing during action control / execution, which in turn allows error monitoring and proactive control in order to avoid error.

In fact, sequential action production can be thought of as the inverse process of event perception. In the former case, a desired goal or regularity is brought about through the performance of a sequence of basic acts (such as picking up objects and manipulating them in various object-specific ways, etc.), while in the latter an observed series of scenes results (somehow) in inferences concerning the goal or culmination of the event and the causal relations between the scenes. Indeed the abductive inference, from the perception of a scene to identification of the event schema to which it (most likely) belongs, is necessary if that event schema is to support scene prediction and event segmentation.

In this paper we argue on the basis of the above parallels for overlap in the cognitive mechanisms that support event perception and action production. We focus first on the domain of action, and one well-developed schema-based account of the control of action, namely the dual-systems "Contention Scheduling / Supervisory System" account of Norman and Shallice (1986). We describe the theoretical account in moderate detail, highlighting the multiple roles of prediction within the theory. In particular, we a) review evidence for the use of prediction within the action control domain, b) demonstrate how predictions can be used to detect and even pre-empt action errors, and c) suggest how prediction, in the form of a forward model of our actions and intentions, fits within the broader cognitive architecture. The implication of this argument is that the abstract representation of consequences is an integral part of the mental representation of action knowledge, and that this applies at multiple levels (i.e., for complex action schemata comprising sequences of subschemata as well as for single basic acts). Extrapolating this to the perception of events, it is argued that event schemata play a parallel role to action schemata, with perceptual experience and current intentions triggering or activating action schemata (in the production of goal-directed sequential actions) and event schemata (in the perception of experiences), and with subsequent processing (including prediction and monitoring) in each domain being driven by those schemata that are most active. The work addresses the question of how, at a mechanistic level, event perception might occur, namely, by essentially running the Contention Scheduling system in reverse.

2 | A SCHEMA-BASED THEORY OF ACTION

2.1 | Contention Scheduling and the Supervisory System

In an influential theoretical account of the organisation and control of sequential action, Norman and Shallice (1986) argued that routine or over-learned behaviour was the product of an automatic schema-based system. Within that system, hierarchically organised action schemata were held to compete for selection and hence control of behaviour through processes of interactive activation. Norman and Shallice (1986) referred to this system for the selection of routine action schemata as Contention Scheduling (CS), and further argued that in situations where deliberate or intentional control was exercised (e.g., non-routine situations, troubleshooting, etc.), the operation of Contention Scheduling could be modulated by a second system, the Supervisory Attentional System (now commonly referred to as the Supervisory System; SS). This second system was held to operate on behaviour only indirectly, by selective excitation, or biasing, of schema representations within CS.

The Supervisory System is held to come into play in situations where no suitable schema exists within CS or where deliberate attention to action is required. Through processes such as planning and strategy formation, SS attempts to generate ordered sequences of known schemata so as to produce an intended outcome, temporarily maintains such ordered sequences, and then selectively excites the schemata within CS in sequence so as to effect the plan or strategy and produce an intended outcome. While the operation of these supervisory processes has not be addressed in detail, SS is assumed to be the source of all intentional activation within the combined CS / SS architecture. Thus, during a routine or everyday task, excitation of the appropriate highest-level node in CS is held to originate in SS. SS is also held to be involved in monitoring the progress of on-going action towards a goal and the inhibition of selected schemata when a mismatch between intention and action is detected (see Shallice and Burgess, 1996).

Support for the above theoretical account of action control comes from two primary sources: slips and lapses in everyday or routine behaviour (as documented in the diary studies of Reason, 1979, 1984; Norman, 1981); and errors in action (particularly goal-directed action) of neurological patients (see citations below). For example, the schema-based nature of action is suggested by *capture errors*, where an intended action sequence is “captured” by an unintended but over-learned routine (e.g., Norman, 1981, p. 8, cites a participant who reported that, after playing cards and then counting pages coming out of a photocopier, “found myself counting 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack, Queen, King”), as well as by the so-called *utilisation behaviour* of certain neurological patients (e.g., Lhermitte, 1983; Shallice et al., 1989; Boccardi et al., 2002), who appear unable to inhibit object-appropriate actions (such as pouring water from a jug into a mug and drinking from the mug) when the objects are available in the immediate environment, even when explicitly instructed not to use those objects.

A wealth of other data, relating primarily to slips and lapses in the routine behaviour of neurologically healthy individuals (Reason, 1979, 1984; Norman, 1981), the more flagrant action errors of various groups of neurological patients (including frontal patients: Luria 1966; Duncan 1986; Schwartz et al. 1991, 1998; left temporo-parietal patients: De Renzi and Lucchelli 1988; Rumiati et al. 2001; and dementia patients: Giovannetti et al. 2002), and action deficits relating to the rate of production in specific populations (Parkinson's Disease patients: Jankovic 2008; and those affected by drug-related amphetamine psychosis: Robbins and Sahakian 1983) also supports the theoretical account.

2.2 | A Computational Model of Contention Scheduling

Cooper and Shallice (2000) developed the verbal account of the control of routine sequential action of Norman and Shallice (1986) into a detailed computational model and demonstrated its potential within the relatively simple task of preparing a cup of instant coffee — a 12 step task based on one component of the breakfast routine of the neurological inpatients of Schwartz et al. (1991). The model was further developed by Cooper et al. (2005) to address a more complex everyday task used by Schwartz et al. (1998) — preparing a packed lunch, comprising a sandwich, drink and snack — and it is this latter implementation that we briefly describe here. Within this model, action schemata are complex structures having the following components:

A goal, which is a state of the environment that will normally be achieved if the schema is executed.

A triggering condition, which specifies the conditions under which the representation of the perceived environment excites the schema.

A (possibly empty) set of subgoal specifications, each element of which is a triple consisting of a goal, a pre-condition that must hold before attempting to achieve the goal, and a post-condition that should hold after successful completion of the goal.

A (possibly empty) set of argument specifications, which specify argument roles for the schema and constraints (in the form of features) on the kinds of things that might fill those roles.

An activation level, which is a scalar value ranging from zero to one. In the absence of any excitatory or inhibitory input, activations tend to a resting level (of 0.1), but excitation may drive this towards the upper limit while inhibition may drive it toward the lower limit.

A status, which is either *selected* or *unselected*. (Norman and Shallice, 1986, also allow schemata to be *dormant*. Dormant schemata are assumed to be effectively "swapped out" of the schema selection process. They have no associated activation level and hence do not to contribute to competitive selection processes, and are ignored here.)

Following Norman and Shallice (1986), the model postulates that action schemata exist at multiple levels within a hierarchy, and that when an action schema is activated above a threshold it is selected. Action schemata at the lowest level, i.e., those with no subgoals, correspond to individual basic actions (e.g., picking up or putting down an object, unscrewing a container, etc.). Selection of a schema at this level is assumed to result in performance of the corresponding basic action. At higher levels, schemata correspond to sets of actions (or more accurately, subgoals). Selection of a higher level schema results in excitation of those subschemata (i.e., those schemata that may achieve one of the subgoals of the higher level schema) whose pre-conditions are met and whose post-conditions are not. Activation therefore effectively flows in a top-down fashion from high-level schemata to low-level schemata. It is assumed that the ultimate source of this activation is the Supervisory System which is held to activate the schema (or schemata) corresponding to the current high-level intention (e.g., to prepare a cup of instant coffee).

Schemata also receive "bottom-up" excitation from the representation of the environment, with the degree of that excitation depending on the salience of schema-relevant objects. Thus, a highly salient object which is within reach and of a size suitable for picking up will excite the "pick-up" schema. These links between the schemata and object representations are bi-directional. Consequently they may create positive feedback loops between the two domains, such that schemata excite appropriate object representations and object representations in turn excite object-appropriate schemata.

If a schema has associated argument specifications, then when the schema is selected each of its arguments is bound to the most salient object representation that meets the argument's specification. Under normal functioning, this and the positive feedback loops between schemata and object representations ensure that schemata and object representations work in synchrony, but if the balance between top-down and bottom-up excitation is disturbed it is possible for object-appropriate schemata to be selected in the absence of top-down intentional control (resulting in the kind of utilisation behaviour produced by some neurological patients).

Within the model the activations of schemata and object representations are updated in a cyclic fashion, with all activations also subject to a lateral inhibitory influence from competing entities (i.e., within the schema network, alternative schemata that might be attempted at the current point in time, and within the object representation network alternative object representations that might fulfil a schema's argument specification) and a self-excitatory influence which partially counter-balances this lateral inhibition. Normally distributed random noise is also assumed to affect all activations on each processing cycle. These three influences — lateral inhibition, self excitation, and noise — implement relatively standard processes of interactive activation and competition (McClelland, 1993), which ensure that, at any point in time, no two competing items can be highly active, and moreover that with sufficient excitation from top-down or bottom-up sources, exactly one item from each subset of competitors will tend to become highly active.

The model as described thus far is able to select a high-level schema, a schema for one of its subgoals, and so on to the level of a basic action, and to apply that action with appropriate arguments. Pre-conditions and post-conditions on a schema's subgoals ensure actions are performed in an appropriate order. Thus, a schema for adding and stirring in an ingredient during beverage preparation will include adding and stirring as separate subgoals, with a pre-condition on stirring being that the ingredient has been added. Finally, selected schemata are inhibited when all post-conditions of all of their subgoals are true. This inhibition results in deactivation (and subsequent deselection) of the selected schema, allowing another schema to become selected, thus yielding a stream of action over time.

The complete model is able to generate appropriate sequences of actions for relatively complex behaviours (e.g., the lunch packing task of Schwartz et al., 1998). Flexibility in behaviour is supported by the use of pre-conditions and post-conditions for regulating the flow of excitation within the schema network (which means that the order of selection of subschemata is not fixed), while opportunistic behaviour (i.e., performing necessary actions when environmentally appropriate) is supported by the bottom-up excitation of schemata. Moreover, when noise in the interactive activation networks is relatively high the model generates the kinds of slips and lapses seen in the behaviour of neurologically healthy individuals when distracted, as described earlier, while with even greater noise or other specific disturbances to activation flow (e.g., between networks) the model is able to reproduce various disorders of action selection (Cooper and Shallice, 2000; Cooper et al., 2005; Cooper, 2007).

3 | PREDICTION AND INTERNAL MODELS IN THE CONTROL OF ACTION

Thus far we have not considered the multiple roles of prediction either within the control of action generally or within the CS / SS theory more specifically. In general terms, and starting from the lowest (temporally least extended) level, prediction has been argued to be essential for the control of rapid sequential action, where motor commands compatible with future states must be generated in advance by the cognitive system (e.g. Wolpert and Ghahramani, 2000).

At a slightly higher level, within CS / SS theory the post-conditions of a schema's subschemata are predictions of the state of the local environment after completion of that subschema. As argued below, these predictions have the potential to support error detection and to thereby trigger error recovery processes. A third role of prediction relates to goal representations associated within schemata. A schema's goal is a representation of both a desired state of affairs associated with the schema and an explicit prediction of the consequences of successful performance of the schema. Higher-level planning processes, assumed to be within the realm of SS, operate over such goal representations in order to generate temporary schemata for the control of non-routine behaviours.

The CS / SS theory provides a framework for formalising these three roles of prediction within sequential action selection and for subsequently specifying them in mechanistic terms. In this section we focus specifically on schema post-conditions and how those post-conditions define a "forward model" — a model of anticipated action consequences that may serve to mitigate error. In the following section we consider how, in an inverted form of the contention scheduling model, such forward models might support perceptual processes of event segmentation.

3.1 | Internal Models in Motor Control

Consider the coordination of rapid fine-grained sequential motor behaviour, such as when serving a tennis ball, playing a musical instrument, or even when typing quickly at a keyboard. These skilled behaviours require that motor commands for upcoming movements be programmed before the preceding movement is complete. Wolpert and Ghahramani (2000) argue that in such cases, the state of the motor system when a subsequent motor command is to be performed must be estimated or predicted, and that this is achieved through the use of several "internal models". Specifically, it is argued that the motor system makes use of:

- a "forward dynamic model", to predict the expected state of the external environment, on the basis of the previous state, the current motor command, and contextual information;
- a "forward sensory model", to predict the expected sensory feedback of a motor command, on the basis of the expected state, the actual motor command, and the context; and
- an "inverse model", to predict the motor command required to achieve a given state, given the current task and context.

These models are held to underlie all motor control (i.e., not just fast fine-grained motor control). Moreover the models are held to feed into each other, with the output of the forward dynamic model feeding into the forward sensory model, the output of the forward sensory model feeding into the inverse model, and the output of the inverse model feeding into the forward dynamic model. The arguments for this sequential cyclic operation are arguably largely aesthetic, however, and an alternative arrangement where the forward models operate in parallel would also seem feasible (i.e., where the state, motor command, and context generate both the anticipated state and sensory feedback in parallel).

To illustrate the three models, Wolpert and Ghahramani (2000) give the example of picking up a milk carton that is thought to be full (and hence heavy). The inverse model allows programming of appropriate reach, grasp and lift motor commands, while the forward dynamic model predicts the consequent state and the forward sensory model predicts the anticipated sensory feedback given that state. If the carton happens to be empty, then this will be apparent from error between the output of the forward models and sensory (visual and proprioceptive) feedback — i.e., the prediction error — which may be used for the on-line adjustment of motor behaviour.

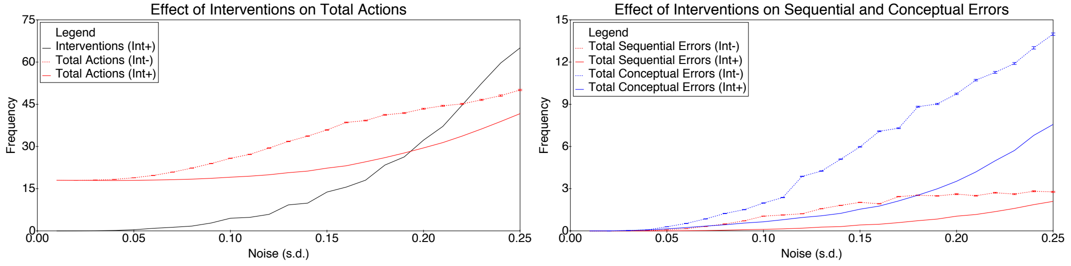

FIGURE 1 The effect of monitoring and intervention based on prediction failure on the errors produced by the Contention Scheduling model of Cooper et al. (2005) when completing a battery of five multiple object tests.

3.2 | Post-Conditions and Forward Models in Contention Scheduling

CS operates at a temporal level above that considered by Wolpert and Ghahramani (2000), and as such the scheduling of schemata does not directly rely on internal models. However, the post-conditions of a schema's subgoals may support monitoring of action outcomes and hence error detection and subsequent correction in a way analogous to motor control. Monitoring based on the checking of post-conditions was included in the implementation of CS described by Cooper et al. (2005), but a systematic investigation of its effects was not reported. Therefore, in order to investigate the role of prediction, post-conditions and monitoring, we here report the results of a novel set of simulations. The model of Cooper et al. (2005) was run on a battery of five multiple object tasks that have previously been used to evaluate the model's behaviour following damage (see Cooper, 2007) with the monitoring/error correction mechanism enabled or disabled (1000 times in each condition), as the level (i.e., standard deviation) of global noise was increased from 0.01 to 0.25. Global noise has previously been argued to simulate a generalised deficit in sequential goal-directed action selection as seen in Action Disorganisation Syndrome. Means and standard deviations were obtained for the number of interventions (in the monitoring / error correction condition) and both the numbers of actions and the numbers of each type of error at each level of noise (in both conditions). Figure 1 provides a comparison of the results across the two conditions.

As shown in the figure, when the error monitoring mechanism is disabled (the "Int−" conditions in the figure), the model produces errors when noise is greater than approximately 0.05. The number of errors (and the number of actions performed) on each task within the battery increases steadily as the level of noise increases. When error monitoring and correction is enabled (the "Int+" conditions in the figure), interventions occur and act to limit errors. Errors still occur, particularly at high levels of noise, but at much lower rates. At moderate levels of noise (between 0.10 and 0.20) the mechanism suppresses all types of sequence errors (i.e., errors involving the misordering of lower-level actions). Within this range, the basic model is also susceptible to so-called conceptual errors (e.g., misusing objects or performing actions at the wrong location). Again, the mechanism for intervening on post-condition failure suppresses these errors, though to a lesser extent than for sequential errors.

It must be emphasised that the intervention mechanism operates only after prediction failure occurs. The recovery mechanism, which in this implementation does little more than try the failed schema again, effectively puts the system back on track, and it is this that prevents error (i.e., the errors produced by the model without interventions are produced subsequent to post-condition failure, they do not reflect behaviour occurring during the faulty execution of a sub-schema leading to that failure). Prediction within the model thus serves as an effective mechanism for detecting situations that will, without supervisory intervention, result in error, thereby signalling the need for proactive control.

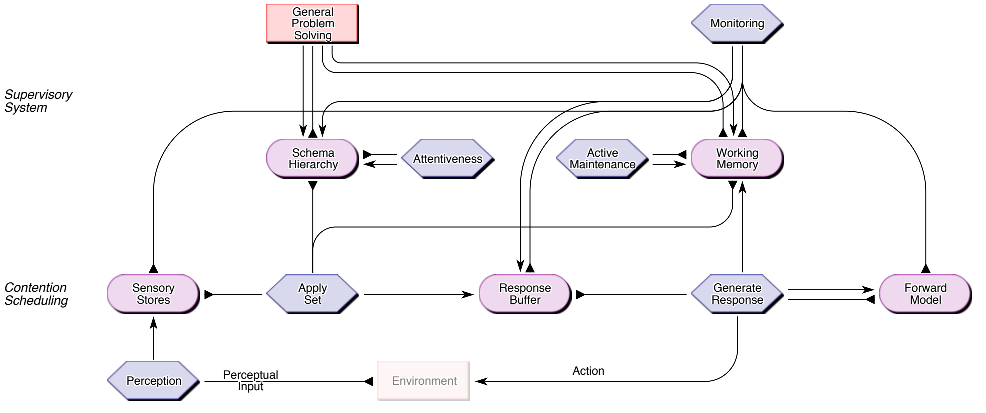

FIGURE 2 The Executive SubProcess (ESPro) framework, adapted from Sood and Cooper (2013). Oblong shaped boxes represent information storage subsystems (buffers). Hexagonal boxes represent process that operate on information stored in buffers (e.g., by transforming it or reactivating it). The rectangular box is a complex process assumed to have internal structure. Blunt-ended arrows represent information read relations between processes and buffers. Sharp-ended arrows represent information write or send operations that might originate from processes. All processes are assumed to function in parallel.

3.3 | Forward Models and Prediction in the ESPro Framework

One question arising from the view of action control as presented here concerns the placement of monitoring within the broader CS / SS theory. One possibility is provided by the Executive SubProcess (ESPro: Sexton and Cooper, 2014; Sood and Cooper, 2013) framework depicted in Figure 2. The figure shows how key functional cognitive components of the CS / SS theory might interact in the generation and regulation of action (and sequential cognition more generally).

Consider first the operation of those components that make up Contention Scheduling. Within the framework, perceptual input is transformed, via the process of *Perception*, into a representation that enters *Sensory Stores*. This representation is assumed to encode the objects available in the immediate perceptual environment and their action-relevant features. If a task-set, or more specifically a schema for mapping sensory inputs to potential responses, is selected (within *Schema Hierarchy*), then *Apply Set* uses this to generate a potential response which is copied or stored in *Response Buffer*. A further process, *Generate Response*, executes the corresponding motor commands. At the same time it updates *Forward Model*, which contains a representation of the anticipated perceptual state(s). Within the framework, schemata (with their activation values) are contained within *Schema Hierarchy*, and *Attentiveness* works to maintain schema activation and regulate competition.

Supervisory processes operate to modulate the functioning of those components described in the previous paragraph. For example, if no schema is sufficiently active so as to exceed the schema selection threshold, then *General Problem Solving* is assumed to apply problem solving heuristics to generate a plan or strategy for the current goal (which is assumed to be accessible within *Working Memory*), and to produce top-down excitation of an appropriate schema within *Schema Hierarchy* (leading to selection of that schema). At the same time, *Monitoring* compares the contents of *Sensory Stores* and *Forward Model*, inhibiting the currently selected schema (in *Schema Hierarchy*) if the two are found to be inconsistent, thereby forcing schema deselection and triggering *General Problem Solving*.

At the time of writing ESPro is more of a generalised framework than a fully-functioning architecture. It has been applied to some standard tasks from the executive function literature, including the Wisconsin Card Sorting Test (Sood and Cooper, 2013) and random number generation (Sexton and Cooper, 2014), both of which tax putative functions such as task setting, response inhibition, and working memory updating, but modelling those tasks has not taxed the *Forward Model* component. Nevertheless, it provides a sketch of the relation between the *Forward Model* and other components of the cognitive system. Thus, with respect to action schemata as conceptualised above, when a schema (at any level) is applied (by *Apply Set*), *Response Buffer* should contain not just the response to be generated but also the schema's post-conditions, which are applied by the *Generate Response* process to update the *Forward Model*.

The proposed location of the *Forward Model* within the framework reflects in part the claim that the model operates with higher level content than the forward models of Wolpert and Ghahramani (2000) proposed within the motor control literature. That is, it contains predicted or anticipated consequences of both lower-level schemata (corresponding to the basic actions described in section 2.2), as well as those of their superordinate schemata (at increasing levels of temporal grain). The justification for this higher level content with the forward model comes in part from the robustness that it provides to the action selection system, as shown by the simulations in section 3.2, and in part from the phenomenology of slip detection (as evidenced by diary reports of catching slips even before they occur, e.g., Norman, 1981, p 12).

Moreover the hierarchical nature of schemata means that the *Forward Model* does not just project one action step forward in time. That is, it does not simply represent the post-conditions of the current low-level schema, but also the post-conditions of its super-ordinate schemata (and all the way up the schema hierarchy). Thus, the predictions contained within the *Forward Model* must include both perceptual / sensory and temporal specifications. *Monitoring* cannot just compare immediate perceptual / sensory input with anticipated perceptual / sensory input, it must also compare the post-conditions of all selected subschemata with anticipations from their ancestors at the given time. This kind of projecting forward in time is needed to catch errors in the act or before they occur and to allow, in the prescient words of (Luria, 1966, p. 238), "the process of comparison of intention and effect".

3.4 | Three Further Roles of Prediction in the Control of Action

While space limitations prevent a full treatment, it should be noted that there are additional potential roles of prediction in action control that have not been considered here. One concerns the setting of cognitive control parameters (e.g., within a theory such as the Executive Control Theory of Visual Attention; Logan and Gordon, 2001) when switching tasks. Consider a naïve participant approaching a colour naming Stroop trial, for example. Having had the task explained, the participant must engage sufficient attention to the colour naming task (which in current models, e.g., Cohen et al., 1990, amounts to setting an appropriate bias to the colour naming pathway). This requires predicting how much attentional bias to colour naming will be sufficient to over-ride the word reading response. Typically, participants will err on the first trial. On this account such errors are due to a poor prediction based on insufficient experience with Stroop stimuli. This prediction is rapidly adjusted following experience such that participants are quickly able to perform the colour naming task without error, provided that strong attentional bias to the task is maintained.

A second role of prediction is evident in the Goal Circuit model of Cooper et al. (2014), which combines aspects of the action control models of Botvinick and Plaut (2004) and Cooper and Shallice (2000). This model uses a simple recurrent network to predict subsequent subgoals within the currently selected schema. As in the work described above, this opens the possibility for anticipatory error correction (when predicted subgoals are not intended subgoals). Such subgoals may also be modulated by higher order control (i.e., the SS) before being fed back into the model (through the goal circuit), making the model capable of using known schemata to perform novel task variants (e.g., preparing coffee with extra measures of sugar).

Finally, prediction has an important role in higher-level planning. While we assume that such planning is the product of *General Problem Solving*, it would equally appear to involve projecting the current state forward in time based on application of a sequence of potential schemata. That is, it would appear to also involve construction of a forward model, but one representing a hypothetical state. Hence, this projection of the current state is presumably deliberately represented in *Working Memory* rather than automatically represented within *Forward Model*, as is required to enable *Monitoring* during the execution of routine behaviour. Note also that this account assumes that *General Problem Solving* is able to reason over the abstract goal representations associated with those schemata known to Contention Scheduling.

4 | EVENT PERCEPTION AS ROUTINE SEQUENTIAL PROCESSING

We began by arguing for similarities between event perception and routine sequential action production. The discussion of the previous two sections – of the Contention Scheduling model of routine sequential action production and the roles of prediction within that (and related) models – might have seemed a detour. However, the Contention Scheduling model is not just a model of routine sequential action production. Rather, it is a model of hierarchical schema-based sequential processing, where schemata can be triggered or activated by both intentional / top-down and environmental / bottom-up sources, and as argued, event perception is guided by hierarchically structured event schemata.

Figure 3 outlines how activation, selection and prediction in the model operate in the case of action schemata (left) and how it might be applied in the case of event perception (right). The outline focusses on the most critical aspects of the model for the two cases. Note though that bottom-up excitation of schemata (as depicted in the event perception model) has always been a feature of the action selection model, where it plays an important role in the explanation of capture errors and utilisation behaviour. Thus, it is not an additional component of the model, though the balance between bottom-up and top-down excitation might be tilted more toward bottom-up sources in the case of event perception and top-down sources in the case of action selection. Note also that in the case of event perception, multiple lower-level scenes may be required to uniquely identify a single event schema at each superordinate level. Thus, from a mathematical perspective, bottom-up excitation might be conceived of as determining a probability distribution over higher-level event schemata. (Similarly, top-down excitation of an action schema can be conceived of as determining a probability distribution over lower-level actions.)

This account of event processing clarifies the sense in which event perception and action production are inverse processes. It suggests that while action production is (generally) the product of top-down, intentionally driven, excitation within a hierarchically organised schema network, event perception is the product of bottom-up, environmentally (or perceptually) driven, excitation within the same network. Both processes generate forward models, which constitute predictions of future states, and both processes are subject to bias or error due to excitation working against the desired flow – bottom-up excitation of action schemata leads to potential utilisation (and other) errors, while top-down excitation of event schemata leads to potential mis-interpretation of events.

Moreover, the account has much in common with several other views that emphasise the roles of action, sensorimotor contingencies, and ideomotor principles in cognition. In the Theory of Event Coding (see in particular Hommel, 2004, 2015), for example, a common code (so-called *event codes*) represents both perceptual events and planned actions. Related accounts have been proposed by O'Regan and Noë (2001), Cisek and Kalaska (2010) and Fuster

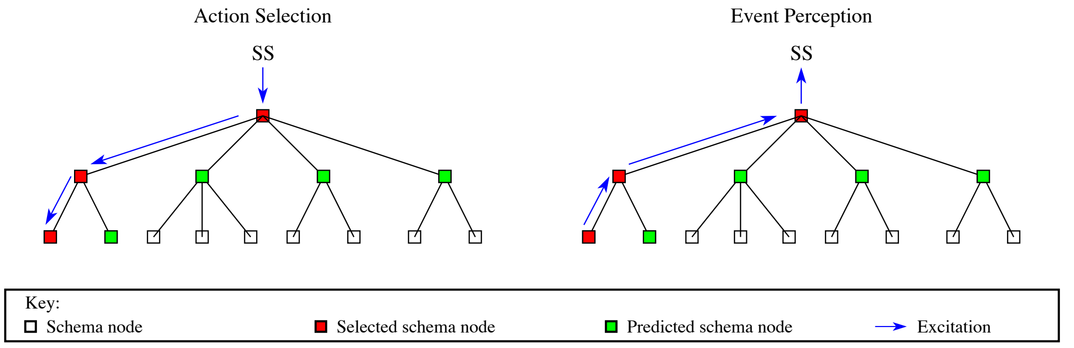

FIGURE 3 Schematic representation of activation flow, schema selection, and prediction in (left) the Contention Scheduling model of routine action selection and (right) the proposed application of the model to event perception. All schema nodes also receive bottom-up excitation in proportion to the degree to which their preconditions match the current state of the perceived environment.

and Bressler (2012). The proposal is also consistent with evidence provided by Flanagan and Johansson (2003), who showed that when participants observed an agent performing a simple sequential task, the participants' gaze was predictive of the next action, rather than reactive (i.e., following the current action). Predictive gaze is a known characteristic of routine sequential action, where visual attention is often focussed on the next action, rather than the current one (Land et al., 1999), suggesting that action execution and event perception (as required in the case of observation) share fundamental processes or representations. This evidence is moreover consistent with both the mirror neuron literature (e.g., Gallese and Goldman, 1998) and studies of imitation and action observation (e.g., Brass et al., 2001; Hamilton and Grafton, 2006), which suggest that observing others act might activate the neural and cognitive representations of actions and their goals.

The above is not to deny that in some cases there are clear differences between actions and events (e.g., in perspective: typically third person/observer versus necessarily first person/actor; in scope: events are more general in that we perceive events that do not have or require an actor; and in the role of agents: which for action is implicit but which for those events with an actor must be explicit). These differences warrant further consideration, and modifications to the existing model may be required to accommodate them.

5 | CONCLUSION

We have reviewed the role of prediction in the generation of goal-directed sequential action, and argued that it may serve roles at a range of levels, from setting cognitive control parameters to allowing monitoring of temporally extended schemata and the detection and capture of likely action errors before they are even produced. Insight into the mechanisms of event perception came from the Contention Scheduling / Supervisory System account of the control of action. The CS / SS account suggests that event schemata may be no different in kind from action schemata. They may simply be schemata that allow us to interpret sensory input. Thus, while action schemata may result in the generation of motor commands and higher-level action plans, event schemata may result in the generation of working memory or otherwise cognitive representations of the sensory input.

Acknowledgements

I am grateful to Martin Butz, four anonymous reviewers, and the participants of the symposium on Event Predictive Cognition at the 2018 Annual Conference of the Cognitive Science Society for constructive criticism which has helped shape this article.

REFERENCES

Arbib, M. A. (1992) Schema theory. *The Encyclopedia of Artificial Intelligence*, **2**, 1427–1443.

Bartlett, F. C. (1932) *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Boccardi, E., Della Sala, S., Motto, C. and Spinnler, H. (2002) Utilisation behaviour consequent to bilateral SMA softening. *Cortex*, **38**, 289–308.

Botvinick, M. and Plaut, D. C. (2004) Doing without schema hierarchies: a recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, **111**, 395–429.

Brass, M., Bekkering, H. and Prinz, W. (2001) Movement observation affects movement execution in a simple response task. *Acta Psychologica*, **106**, 3–22.

Cisek, P. and Kalaska, J. F. (2010) Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, **33**, 269–298.

Cohen, J. D., Dunbar, K. and McClelland, J. L. (1990) On the control of automatic processes: a parallel distributed processing account of the stroop effect. *Psychological Review*, **97**, 332–361.

Cooper, R. P. (2007) Tool use and related errors in ideational apraxia: The quantitative simulation of patient error profiles. *Cortex*, **43**, 319–337.

— (2016) Schema theory and neuropsychology. In *From Neuron to Cognition via Computational Neuroscience* (eds. M. A. Arbib and J. J. Bonaiuto), chap. 14, 433–456. The MIT Press.

Cooper, R. P., Ruh, N. and Mareschal, D. (2014) The goal circuit model: A multi-route model of the acquisition and control of sequential action in humans. *Cognitive Science*, **38**, 244–274.

Cooper, R. P., Schwartz, M. F., Yule, P. and Shallice, T. (2005) The simulation of action disorganisation in complex activities of daily living. *Cognitive Neuropsychology*, **22**, 959–1004.

Cooper, R. P. and Shallice, T. (2000) Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, **17**, 297–338.

De Renzi, E. and Lucchelli, F. (1988) Ideational apraxia. *Brain*, **111**, 1173–1185.

Duncan, J. (1986) Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, **3**, 271–290.

Flanagan, J. R. and Johansson, R. S. (2003) Action plans used in action observation. *Nature*, **424**, 769–771.

Fuster, J. M. and Bressler, S. L. (2012) Cognit activation: a mechanism enabling temporal integration in working memory. *Trends in Cognitive Sciences*, **16**, 207–218.

Gallese, V. and Goldman, A. (1998) Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, **2**, 493–501.

Giovannetti, T., Libon, D. J., Buxbaum, L. J. and Schwartz, M. F. (2002) Naturalistic action impairments in dementia. *Neuropsychologia*, **40**, 1220–1232.

Hamilton, A. F. d. C. and Grafton, S. T. (2006) Goal representation in human anterior intraparietal sulcus. *Journal of Neuroscience*, **26**, 1133–1137.

Hommel, B. (2004) Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, **8**, 494–500.

— (2015) The theory of event coding (tec) as embodied-cognition framework. *Frontiers in Psychology*, **6**, 1318.

Jankovic, J. (2008) Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry*, **79**, 368–376.

Kurby, C. A. and Zacks, J. M. (2008) Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, **12**, 72–79.

Land, M., Mennie, N. and Rusted, J. (1999) The roles of vision and eye movements in the control of activities of daily living. *Perception*, **28**, 1311–1328.

Lhermitte, F. (1983) 'Utilization behaviour' and its relation to lesions of the frontal lobes. *Brain*, **106**, 237–255.

Logan, G. D. and Gordon, R. D. (2001) Executive control of visual attention in dual-task situations. *Psychological Review*, **108**, 393–434.

Luria, A. (1966) *Higher Cortical Functions in Man.* London: Tavistock.

McClelland, J. L. (1993) Toward a theory of information processing in graded, random, and interactive networks. In *Attention and Performance, XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience* (eds. D. E. Meyer and S. Kornblum), 655–688. Cambridge, MA: The MIT Press.

Norman, D. A. (1981) Categorization of action slips. *Psychological Review*, **88**, 1–15.

Norman, D. A. and Shallice, T. (1986) Attention to action: Willed and automatic control of behaviour. In *Consciousness and Self-Regulation* (eds. R. Davidson, G. Schwartz and D. Shapiro), vol. 4, 1–18. Springer.

O'Regan, J. K. and Noë, A. (2001) A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, **24**, 939–973.

Reason, J. T. (1979) Actions not as planned: The price of automatization. In *Aspects of Consciousness, Volume 1: Psychological Issues* (eds. G. Underwood and R. Stevens), chap. 4, 67–89. London, UK: Academic Press.

— (1984) Lapses of attention in everyday life. In *Varieties of Attention* (eds. W. Parasuraman and R. Davies), chap. 14, 515–549. Orlando, FL: Academic Press.

Robbins, T. W. and Sahakian, B. J. (1983) Behavioral effects of psychomotor stimulant drugs: clinical and neuropsychological implications. In *Stimulants: Neurochemical, Behavioral, and Clinical Perspectives* (ed. I. Creese), 301–338. New York, NY: Raven Press.

Rumiati, R. I., Zanini, S., Vorano, L. and Shallice, T. (2001) A form of ideational apraxia as a delective deficit of contention scheduling. *Cognitive Neuropsychology*, **18**, 617–642.

Schmidt, R. A. (1975) A schema theory of discrete motor skill learning. *Psychological Review*, **82**, 225–260.

Schwartz, M. F., Montgomery, M. W., Buxbaum, L. J., Lee, S. S., Carew, T. G., Coslett, H., Ferraro, M., Fitzpatrick-DeSalme, E., Hart, T. and Mayer, N. (1998) Naturalistic action impairment in closed head injury. *Neuropsychology*, **12**, 13.

Schwartz, M. F., Reed, E. S., Montgomery, M., Palmer, C. and Mayer, N. H. (1991) The quantitative description of action disorganisation after brain damage: A case study. *Cognitive Neuropsychology*, **8**, 381–414.

Sexton, N. J. and Cooper, R. P. (2014) An architecturally constrained model of random number generation and its application to modeling the effect of generation rate. *Frontiers in Psychology*, **5**, 1–14.

Shallice, T. and Burgess, P. W. (1996) The domain of supervisory processes and temporal organization of behaviour. *Philosophical Transactions of the Royal Society of London*, **B351**, 1405–1412.

Shallice, T., Burgess, P. W., Schon, F. and Baxter, D. M. (1989) The origins of utilization behaviour. *Brain*, **112**, 1587–1598.

Sood, M. and Cooper, R. P. (2013) Modelling the supervisory system and frontal dysfunction: An architecturally grounded model of the wisconsin card sorting task. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (eds. M. Knauff, M. Pauen, N. Sebanz and I. Wachsmuth), 1354–1359. Berlin, Germany: Cognitive Science Society Incorporated.

Wolpert, D. M. and Ghahramani, Z. (2000) Computational principles of movement neuroscience. *Nature Neuroscience*, **3**, 1212.

Zacks, J. M. and Swallow, K. M. (2007) Event segmentation. *Current Directions in Psychological Science*, **16**, 80–84.

Zacks, J. M. and Tversky, B. (2001) Event structure in perception and conception. *Psychological Bulletin*, **127**, 3–21.