TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 13 (2021) 10–24

ISSN:1756-8765 online
DOI: 10.1111/tops.12522

This article is part of the topic "Forthcoming Topic: Event-Predictive Cognition: From Sensorimotor via Conceptual to Language-Based Structures and Processes," Martin V. Butz, David Bilkey, and Alistair Knott (Topic Editors). For a full listing of topic papers, see https://onlinelibrary.wiley.com/toc/17568765/2021/13/1

Event-Predictive Cognition: A Root for Conceptual Human Thought

Martin V. Butz,[a] Asya Achimova,[a] David Bilkey,[b] Alistair Knott[c]

[a]*Neuro-Cognitive Modeling Group, Department of Computer Science, Department of Psychology, Faculty of Science, University of Tübingen*
[b]*Department of Psychology, University of Otago*
[c]*Department of Computer Science, University of Otago*

Received 24 June 2019; received in revised form 11 October 2020; accepted 11 October 2020

Abstract

Our minds navigate a continuous stream of sensorimotor experiences, selectively compressing them into events. Event-predictive encodings and processing abilities have evolved because they mirror interactions between agents and objects—and the pursuance or avoidance of critical interactions lies at the heart of survival and reproduction. However, it appears that these abilities have evolved not only to pursue live-enhancing events and to avoid threatening events, but also to distinguish food sources, to produce and to use tools, to cooperate, and to communicate. They may have even set the stage for the formation of larger societies and the development of cultural identities. Research on event-predictive cognition investigates how events and conceptualizations thereof are learned, structured, and processed dynamically. It suggests that event-predictive encodings and processes optimally mediate between sensorimotor processes and language. On the one hand, they enable us to perceive and control physical interactions with our world in a highly adaptive, versatile, goal-directed manner. On the other hand, they allow us to coordinate complex social interactions and, in particular, to comprehend and produce language. Event-predictive learning segments sensorimotor experiences into event-predictive encodings. Once first encodings are formed, the mind learns progressively higher order compositional structures, which allow reflecting on the past,

Correspondence should be sent to Martin V. Butz, Department of Computer Science, Faculty of Science, University of Tübingen, Sand 14, Tübingen 72076, Germany. E-mail: martin.butz@uni-tuebingen.de

[Correction added on December 17, 2020, after first online publication: The title was changed from 'Editors' Review and Introduction: Event-Predictive Cognition: From Sensorimotor via Conceptual to Language-Based Structures and Processes' to 'Event-Predictive Cognition: A Root for Conceptual Human Thought'.]

reasoning, and planning on multiple levels of abstraction. We conclude that human conceptual thought may be grounded in the principles of event-predictive cognition constituting its root.

Keywords: Theory of event coding; Event segmentation theory; Action events; Anticipatory behavior; Natural language processing; Bayesian brain; Predictive coding; Cognitive ontogeny

1. Introduction

Our behavior and our thoughts develop from our embodied sensorimotor experiences. To understand the human mind and intelligence, it is thus necessary to understand which learning processes, selective activation mechanisms, and concurrently developing encodings continuously determine actual behavior and thought. Over the last years, it has become increasingly clear that our brain constitutes a generative processing system, which learns predictive, approximate Bayesian models purposefully to optimize its own behavior for maintaining internal homeostasis (Friston, 2009; Maturana, 1975).

However, mere probabilistic, predictive processing does not seem to be enough. In order to effectively interact with each other and the rest of the world in progressively more complex societies, behavior needs to be chosen and executed in highly flexible, versatile, context-dependent manners. When complex social interactions are to be coordinated, communication and ultimately language comprehension and production become necessary. As a result, compacted, conceptual, and loosely hierarchically structured encodings are needed to enable individual and social compositional reasoning and planning. Research on event-predictive cognition addresses the question of how our brain solves the involved challenges.

More than 30 years ago, Marr (1982) asked the question "What does it mean, to see?," emphasizing that seeing is about the interactions between neural activities (Marr's neural implementation level), algorithmically unfolding processes (Marr's algorithmic level), and cognitive, computational mechanisms (Marr's computational level). It is crucial to bridge these three levels of analysis and to identify critical interactions between them, when the final goal is to understand how our mind works (Butz & Kutter, 2017; Griffiths, Vul, & Sanborn, 2012).

Accordingly, this special issue on event-predictive cognition (EPCog) addresses the neural structures, algorithmic processes, and unfolding computations involved in creating the compact, loosely hierarchically structured encodings needed to coordinate versatile and adaptive, socially interactive behavior. Taken together, the contributions in this volume essentially suggest that event predictions lie at the heart of human thought. Event-predictive processes, encodings, and their (ontogenetic and phylogenetic) development are scrutinized by experts from the following disciplines: neurobiology, cognitive and computational neuroscience, developmental and behavioral psychology, artificial intelligence and machine learning, and linguistics. We hope this collection will thus contribute to the development of an integrative perspective on how our developing brain-body complex yields our human minds, including thoughts, behavior, and intelligence.

2. Event as units of human experience

EPCog as a research topic links the concept of events, including event boundaries and event transitions, with the predictive coding and processing perspective. EPCog sets out to explore the extent to which event-predictive encodings and processes foster the development of abstract, conceptual, compositionally recombinable structures from sensorimotor experiences. It links event-predictive, conceptual structures to both sensorimotor and language structures.

2.1. Event conceptualizations in behavior and language

Event encodings have been proposed at numerous levels of cognitive processing. At rudimentary sensorimotor levels, the theory of event coding (TEC; Hommel, Müsseler, Aschersleben, & Prinz, 2001) has proposed that actions and their effects are compressed into a common code. This common code is assumed to be essential for enabling the invocation of flexible, anticipatory behavior. The ideomotor principle suggests that common codes develop from initially reflex-like behavior (Herbart, 1825; Hoffmann, 1993; Prinz, 1990; Stock & Stock, 2004), in that, for example, a simple arm stretching tendency soon develops into progressively intentional, goal-directed hand reaching behavior. As a result, progressively more complex anticipatory behavior control abilities develop.

At a somewhat more abstract level, it was shown that humans tend to perceive and segment events consistently across participants (Zacks & Tversky, 2001). An event may be generally characterized as “[...] a segment of time at a given location that is conceived by an observer to have a beginning and an end” (Zacks & Tversky, 2001, p. 3). Motivated by various segmentation studies, Zacks, Speer, Swallow, Braver, and Reynolds (2007) proposed the event segmentation theory (EST), which suggests that events are encoded, perceived, and processed in our brain as integrated units of thought. Moreover, EST suggests that we segment our continuous perceptual sensorimotor stream of information into events by analyzing prediction error dynamics, inducing event segmentations when surprising patterns are detected.

Even at the language level, events seem to play a fundamental role. Gärdenfors (2014, p. 107) has gone so far as to suggest that “sentences express events,” whereby multiple types of events can be contrasted. *Stative events* specify states and situations that are non-dynamic, durative, and atelic. *Dynamic events* can be further differentiated into being (a) durative and atelic (i.e., ongoing activity, unbounded), (b) durative and telic (i.e., an activity that is bounded by an accomplishment), or (c) instantaneous and telic (i.e., a direct achievement; Casati & Varzi, 2015; Gärdenfors, 2014; Vendler, 1957). Jackendoff (2002, p. 123) suggests that “thoughts expressed by language are structured in terms of a cognitive organization called *conceptual structure* (CS). Conceptual structure is not part of language per se—it is part of thought” (author’s emphasis). Accordingly, this volume explores to what extent thoughts, and the conceptual structures with which they are formed, may be rooted in event-predictive processes and encodings.

2.2. Event-predictive codes

Besides the importance of event-respective conceptualizations, a close linkage to a rather universal computational encoding principle in our minds, that is, predictive coding, has become apparent (Bar, 2009; Barlow, 1961; Butz, 2008; Butz, Sigaud, & Gérard, 2003a; Friston, 2002; Hohwy, 2013; König & Krüger, 2006; Pezzulo, Butz, Sigaud, & Baldassarre, 2009; Rao & Ballard, 1999). In his seminal paper on modeling language with temporally predictive, recurrent artificial neural networks, Elman (1990, p. 207) focused on "temporal events," concluding that hierarchical structures may emerge from analyzing temporal event sequences and that "the error signal is a good metric of where structure exists." Interestingly, EST uses a closely related error signal as an event boundary indicator.

When adapting this predictive coding stance, events may be viewed as being encoded by a network of active predictive codes. These codes will predict each other's activities, forming distinct attractor states, which minimize mutual prediction error. Concurrently, the active codes will predict particular aspects that characterize the currently unfolding event, including the properties of the involved entities, along with their roles, spatial-relational dynamics, and likely behavior. Intentional and motivational reward-related aspects of the unfolding event may also be predicted. For example, reaching for an object may be encoded by predicting that a hand (or another agent) will move to a target object such that the relative distance will dynamically decrease toward zero. In addition, object movability may be encoded and possibly distinct reward expectations may be activated by particular objects, such as a bottle of water, which would generate the expectation of refreshment and quenching thirst. Moreover, the beginnings and endings of events—that is, *event boundaries*—may be predicted when critical circumstances apply, effectively encoding the conditions under which particular events typically commence or end. For the reaching example, this would correspond to codes that predict that an object is reachable and that the hand starts to move toward it—marking the beginning of a reaching event—or to codes that predict tactile feedback and object movement—marking the end of a reaching event upon which the object is touched and thus affected by the hand.

In the case of a stative event, the involved predictive codes form an attractor that may predict, for example, stable spatial relations, properties of the typically involved entities, or more global spatial and temporal characteristics. For example, when stating that "the ball lies in the suitcase," a predictive network may be instantiated that predicts that a ball entity is stably located in the interior of a suitcase entity (Butz, 2017).

In the case of a dynamic event, the dynamic attractor includes dynamically changing, temporally predictive codes, which may encode changes in state properties (e.g., spatial relations, involved entities, emotions), focusing on the manner (i.e., the dynamics of the motor or force activities), on the change in spatial relations, or on the final result. In other words, a dynamic event will be encoded by particular motor or force dynamics and how these dynamics affect particular environmental and bodily states, which cause the perception of interaction-characteristic sensory signal dynamics.

For example, "climbing" implies a durative, atelic event, which may be encoded by an entity code whose position relative to a slope is predicted to increase upward. However, "she climbed the mountain" describes a durative, telic event, which essentially includes the final event boundary, that is, a prediction that the person reached the mountain top as the final state. Similarly, "she reached the top of the mountain" fully focuses on the event boundary, that is, on an instantaneous and telic event, which may be encoded by predicting that the distance to the mountain top reached zero (Gärdenfors, 2014).

The exact nature of such predictive event codes, and particularly also their actual algorithmic and neural implementation in our brains, certainly remains largely unknown. Nonetheless, as summarized below, the contributions in this volume hint at particular and integrative details on all three of Marr's levels of explanation.

2.3. Event-predictive processing and learning

How can these event-predictive codes develop from basic sensorimotor experiences? Clearly, during ontogenetic development, sensorimotor experiences need to be structured into progressively more complex, compact codes. Starting bottom up with mere self-generated simple actions, very basic sensorimotor experiences can be expected to develop into TEC-like common predictive codes, which predict motor-dependent sensory and deeper perceptual effects. These codes may then be actively explored further, enabling the derivation of progressively more complex codes. It appears that the involved highly interactive processes thus feed on each other: Event-predictive codes develop from the event-predictive analysis of sensorimotor experiences; they are analyzed, segmented, and compacted given the so-far learned encodings, as suggested by EST; moreover, they are actively explored dependent on the so-far learned encodings, as suggested by the ideomotor principle and related theories of anticipatory behavioral control. Over time, progressively more compacted, loosely hierarchically structured event-predictive codes may emerge, which may be closely related to pre-linguistic conceptual structures and event construals (Butz, 2016, 2017; Gärdenfors, 2014; Gumbsch, Otte, & Butz, 2017).

A body of evidence has accumulated that suggests that active exploration and learning may be formalized by free energy minimization (Friston, 2009, 2010), which, informally speaking, determines goal-directed behavior, adapts internal state estimates, and pursues model learning by minimizing (anticipated) prediction error. With respect to decision-making and control, formalizations that take an expected future horizon into account yield curiosity-driven and goal-directed planning, reasoning, and motor behavior. The resulting so-called active inference chooses those (motor and cognitive) activities that are believed to lead to the achievement of desired states while minimizing anticipated surprise (Botvinick & Toussaint, 2012; Friston, FitzGerald, Rigoli, Schwartenbeck, & Pezzulo, 2018). Similarly, learning as well as the adaptation of current state estimations can be formalized as probabilistic (approximately Bayesian) model inference (Friston, 2003, 2009).

Despite its high potential, the free energy principle does not formalize particular ontogenetic learning constraints—or inductive learning biases—that have likely evolved phylogenetically and that guide the development of our brains and minds (Butz, 2017; Butz

& Kutter, 2017; Lake, Ullman, Tenenbaum, & Gershman, 2017). As detailed above, evidence is accumulating that our brain actively attempts to compact our sensorimotor experiences into event-predictive encodings. In light of this evidence, we can speculate that active inference should be augmented with inductive biases that focus neural activities on the dedicated and compact processing of event and event-boundary signals. How exactly such inductive biases are implemented remains to be shown (several options are scrutinized in this volume; cf. Baldwin & Kosie, 2021; Shin & DuBrow, 2021).

Event-predictive encodings with some degree of hierarchical structure are known to be highly useful when complex goal-directed planning and decision-making is needed (Barto & Mahadevan, 2003; Botvinick, Niv, & Barto, 2009). Moreover, when reflecting on the past or imagining possible—or even fully hypothetical—future or counterfactual situations and the unfolding dynamic events within, abstractions away from current sensorimotor perceptions are mandatory (Buckner & Carroll, 2007). Finally, when linking these event-predictive encodings to language, conceptual structures (Jackendoff, 2002) and "event construals" (Gärdenfors, 2014) are necessary to enable semantic sentence processing.

To scrutinize the nature of event-predictive structures, their development, as well as the processes that unfold within these structures, an interdisciplinary, integrative research approach is necessary. This volume thus attempts (a) to link the knowledge available about sensorimotor processing with conceptual encoding levels, and (b) to link these conceptual, pre-linguistic encoding levels with language. From this perspective, this volume essentially contributes to the language grounding challenge and the question of how language readiness and competence may emerge from a focused analysis of sensorimotor experiences.

3. An interdisciplinary perspective on the mind

Psychological studies of development, memory, and behavior all indicate that events play critical roles in cognitive processes. Evidence from neuroscience also corroborates the idea that predictive encodings bind behaviorally critical aspects of real-time experiences into compact, event-predictive encodings. Finally, recent computational modeling efforts suggest that events and event boundaries can be identified accurately by machine learning algorithms. When focusing learning on event boundaries, the resulting event-based machine learning algorithms yield event-based bindings and clusterings (Gumbsch, Butz, & Martius, 2019; Gumbsch, Kneissler, & Butz, 2016; Zacks et al., 2007). These algorithms combine several benefits. They support accounts of event and episodic memory formation, they promise solutions to the catastrophic forgetting problem, and they suggest methods for optimizing behavioral primitives. They also contribute to models of habit formation and to models of hierarchical goal-directed planning and reasoning. Moreover, these sensorimotor-grounded structures appear to be well-suited to be linked to language, offering intermediate, semantic, sensorimotor grounded, conceptual, compositionally recombinable structures (Jackendoff, 2002).

This special issue brings together cognitive scientists who are experts in developmental, cognitive and neuro-computational psychology, linguistics, machine learning, and cognitive and computational neuroscience with the purpose to foster the development of an overarching perspective on event-predictive cognition. Ultimately, these insights may reveal how the human mind develops, which event-predictive structures constitute our knowledge and memory, and which processing dynamics unfold within those structures to generate actual thought and human intelligence.

Accordingly, this special issue addresses the following core questions:

- What are the origins of event-predictive processes and why are they useful for improving behavior?
- How is sensorimotor dynamics selectively integrated into event-predictive encodings?
- Can event-predictive encodings serve as a basis for modeling conceptual, pre-linguistic structures?
- How can such encodings be linked to language?
- Which signals (e.g., prediction errors) best support the development of suitable abstractions, compact encodings, and event-taxonomies?
- How is missing information about events inferred and integrated with available information?
- What happens when event predictions go wrong?

4. Paper contributions and connections

This volume starts with a novel perspective on the phylogenetic development of event processing units, that is, neurons themselves. Paulin and Cahill-Lane (2021) argue that in a particular environmental niche 560 million years ago—just a few million years before the Cambrian explosion—some of the motile animals that fed on microbial mat-grounds started to feed on each other, resulting in a dynamic predator–prey situation. The authors corroborate evidence that the evolution of event-predictive neurons allowed animals to feed as long as possible while avoiding to be eaten since they were now able to flee just in time, that is, when the "being eaten" event onset was imminent. While the resulting neural event processing is only implicitly anticipatory (Butz & Kutter, 2017; Butz, Sigaud, & Gérard, 2003b), optimized by evolutionary process, it may indeed have laid the sensorimotor-grounded basis for all more complex event processing abilities thereafter.

The roots of event-predictive cognition in motor control and action decision-making are then pursued from psychological and cognitive modeling perspectives. Elsner and Adam (2021) argue that the perception and interpretation of events must be at least partially rooted in own-action competencies. Adopting a neural network modeling perspective, Cooper (2021) points out that action execution and event perception are closely linked, in that both processes involve an event representation: In the case of action execution, the event representation is first expressed as a goal, which generates sensorimotor behavior, while in event perception, the event representation is constructed from sensorimotor signals. Elsner and Adam (2021) and Cooper (2021) essentially agree that action goals, decision-making, action production and control, as well as action monitoring are closely linked, and may be subsumed by the notion of event-predictive processing. Interestingly, action monitoring by event-predictive activities directly enables the distinction between self-agency and the agency of others (Elsner & Adam, 2021), potentially causing problems when the event-predictive activities are erroneous (Storchak, Ehlis, & Fallgatter, 2021).

Being thus rooted in the need for event-predictive decision-making and motor control, the increasing versatility of behavior requires that events are appropriately segmented and compactly encoded. The developing event structures may be characterized by predictive encoding attractors, which reliably apply while an event is processed or executed (Baldwin & Kosie, 2021; Butz, 2016). Shin and DuBrow (2021) emphasize the need to infer latent variables, which encode the causes that constitute an event. Event transitions then manifest themselves in significant changes in the activities of latent variables. Baldwin and Kosie (2021) point out that these changes are often predictably unpredictable, which may facilitate the detection of event boundaries in the first place. These authors helpfully identify two possible methods for detecting event boundaries. First, the dedicated processing of surprising sensorimotor information, that is, unexpectedly large errors between the predicted event dynamics and actual perceptions (Baldwin & Kosie, 2021; Cooper, 2021; Stawarczyk, Bezdek, & Zacks, 2021; Ünal, Ji, & Papafragou, 2021) can serve as event boundary indicators. This initially exogenous surprise signal, once integrated into event and event-boundary predictive encodings, becomes endogenously predicted, thus suppressing (the then predictable) surprise after learning (Baldwin & Kosie, 2021; Stawarczyk et al., 2021; Storchak et al., 2021). Second, the fact that event processing involves the inference of latent variables can be used in service of event boundary detection. Latent variables, which may be understood as setting the stage for processing particular events via their predictive activities, must necessarily differ when different events are processed (Shin & DuBrow, 2021). As a result, tracking latent variable activity changes across time may also serve to identify event boundaries. Hohwy, Hebblewhite, and Drummond (2021) highlight the strong linkage between (latent) event encodings and the predictive brain perspective, where Bayesian inference processes should lead to the development of probabilistic, loosely hierarchically structured, event-predictive encodings. Recent predictive, Bayesian information processing and neural network models offer algorithms that implement these two perspectives on event segmentation (Butz, Bilkey, Humaidan, Knott, & Otte, 2019; Gumbsch et al., 2019). Meanwhile, Hebblewhite, Hohwy, and Drummond (2021) emphasize the strong linkage between event encodings and hierarchical reinforcement learning approaches. The authors stress the need to still clarify how to control the granularity of event-predictive segmentations and levels of abstractions thereof. Moreover, they propose that the options framework in hierarchical reinforcement learning (Butz & Kutter, 2017; Sutton, Precup, & Singh, 1999), in combination with option-focused policy gradient techniques (Bacon, Harb, & Precup, 2017), might offer a useful, complementary mechanism for producing behaviorally and cognitively useful abstractions.

The event structures developed from these methods integrate relevant agents and entities, and identify the distinct roles they play in events. Knott and Takac (2021) emphasize that such role structures, while being rooted in actual sensorimotor-based manipulations of the environment, can be closely linked to the hierarchical logical structure of sentences in human language. They suggest a general framework for characterizing sensorimotor processes as sequences of discrete "deictic operations"—and a framework for characterizing event representations that tap into these same deictic operations. Deictic operations include perceptual operations (for instance, attention to, and classification of, an external agent or object) but also motor operations (for instance, the action of reaching toward an attentionally selected item in peripersonal space). Their account of the interface between event representations and language retains a reference to deictic operations: Thus, for Knott and Takac, deictic operations are critical in an account of how language interfaces with sensorimotor processes.

In Knott and Takac's conception, event encodings have a symbolic flavor. In the model of McRae, Brown, and Elman (2021), however, event representations are not symbolic. McRae et al., working with recurrent neural network models, point out that temporally predictive structures may be essential for the characterization of particular events. The predictive mechanisms in this case encode the entities and actions involved in an event, as well as their typical interactions over time. These structures are analogue, rather than symbolic. Moreover, the progression from one event to another is non-linear, enabling flexibility in the imagination and the control of an event. Nonetheless, loose hierarchical structures can still be encoded, which may characterize the essence of more complex events, such as "changing the tire of a car."

Separately from questions about the structure of events, there are also questions to be asked about the content of events. Which entities, circumstances, and other properties should be integrated into an event representation? The inference of the latent causes, which characterize an event, is typically non-trivial (Shin & DuBrow, 2021). Accordingly, Storchak et al. (2021) point out that incorrectly processed error signals may lead to the false integration or non-integration of event signals, potentially leading to inappropriate stimulus interpretations—essentially generating exogenous surprise signals when they should have been predicted endogenously. A related challenge lies in determining whether a currently perceived event is novel, or should be integrated into available event-predictive encodings (Baldwin & Kosie, 2021; Stawarczyk et al., 2021).

While it remains unresolved when sensorimotor signals are integrated into available event-predictive structures and when novel structures are formed (probably again these two mechanisms are blended rather than distinct), many papers emphasize the role of working memory in this process (Bilkey & Jensen, 2021; Knott & Takac, 2021; McRae et al., 2021; Stawarczyk et al., 2021). Stawarczyk et al. (2021) in particular focus on the role of the brain's default network for maintaining and processing particular events and event successions. A push–pull situation is described, which continuously regulates information processing over time, assigning two roles to the default network core. One lies in generating events during imagination, reasoning, and planning, when decoupled from the sensorimotor here and now. Another role lies in integrating incoming sensorimotor information into event-predictive processes, to maintain internal latent activities that encode and process the current event dynamics. The generation of meaningful imaginations relies on the proper re-combination of events and event successions, where "properness" may be measured by the mutual prediction error in event-predictive attractors. Such imaginations essentially enable compositional, and even counterfactual, reasoning, planning, and reflection processes (Bilkey & Jensen, 2021; McRae et al., 2021; Storchak et al., 2021; Ünal et al., 2021). While being in tune with the world, however, event-predictive processes will continuously work on optimally integrating sensorimotor information, adapting event-predictive encoding activities, and thus parsing, perceiving, and generating events, including action events (Cooper, 2021; Stawarczyk et al., 2021). As a result, while processing and executing events, attention tends to focus on those regions that are anticipated to be highly informative in interpreting or controlling current event dynamics (Baldwin & Kosie, 2021; Elsner & Adam, 2021; Knott & Takac, 2021; Paulin & Cahill-Lane, 2021; Stawarczyk et al., 2021).

Besides the important role of a default network, Bilkey and Jensen (2021) as well as Baldwin and Kosie (2021) put forward the likely critical role of the hippocampus for learning event and episodic memory structures. Bilkey and Jensen (2021) view event boundary markers as critical components for the consolidation of event episodes and discuss the neural signatures that might accompany them, linking the hippocampal literature on putative event boundary markers with that describing memory consolidation. They also note their potential importance for effective recombination with other compatible events and event boundaries, enabling the formation of novel event successions. There are indeed hints that this principle may be implicated both in the formation of episodic memories and in the learning of spatial representations, which keep local proximity estimates but generalize over deeper temporal orders. Note that a close relation to temporally predictive artificial neural network structures can be drawn here (Cooper, 2021; Knott & Takac, 2021; McRae et al., 2021), whereby the inference of latent variables may be critical to avoid catastrophic forgetting while still ensuring effective memory consolidation (Kumaran, Hassabis, & McClelland, 2016; Shin & DuBrow, 2021). It thus appears that with the help of the hippocampus, default network core structures can be re-instantiated to consolidate sensorimotor segmentations and event-predictive encodings. In this case, learning progress and the selective, active exploration and consolidation may be viewed as a push–pull process. The world pulls the process toward novel experiences that seem to be integratable or that may enhance the so far available event-predictive latent knowledge structures. Meanwhile, these structures also push cognition toward intentionally optimizing them further, by directing attention and actively exploring the environment (Baldwin & Kosie, 2021; Cooper, 2021; Shin & DuBrow, 2021).

The linkage between event-predictive structures and language is investigated in detail by several of the papers in this volume. Ünal et al. (2021) show that event boundary and event role encodings do indeed align well with linguistic encodings of events. Knott and Takac (2021) show how sensorimotor interaction encodings offer themselves as the basis for grammatical encodings at the logical form level. However, the neural model of McRae et al. (2021) gets as input distributed event encodings, including agent, action, and patient. It then focuses on the loosely hierarchically structured network of subevents, which are constitutive of a more complex event, such as changing a tire.

Warren and Dresang (2021) further summarize these papers and additional data that indicate strong interactions between language and event-predictive structures. Moreover, they highlight that event-predictive interpretations of linguistic expressions may very well stay underspecified and may be made concrete in various manners only when necessary. As Warren and Dresang (2021) emphasize, this observation may explain the variability in describing events observed in the data and model of McRae et al. (2021), which seems to, but does not, stand in contrast to the rather strict hierarchical structures proposed in Cooper (2021) and implied in Knott and Takac (2021). The social, communicative, and cooperative utility of event encodings is emphasized by several papers (Baldwin & Kosie, 2021; Elsner & Adam, 2021; Ünal et al., 2021). When interpreting the actions of others we are able to infer their intentions, essentially anticipating their current physical or communicative goals, enabling us to respond in preemptive, anticipatory manners. The presented results and theoretical considerations suggest that two means of analyses develop, which go hand in hand. On the one hand, the child monitors the semantics of actions and environmental interactions, inferring and anticipating unfolding event dynamics and progressions thereof. On the other hand, the child develops the ability to process aspects of the world in an event-predictive, linguistically driven manner, enabling grammatically constraint inferences and reasoning processes while remaining grounded in the experienced reality. As a result, communication can unfold on a linguistic, event-predictive level, lifting socially interactive behavior onto a much more advanced, versatile, and adaptive stage.

Finally, Kuperberg (2021) offers an encompassing, integrative theoretical treatise of nearly all papers covered in this volume. Relying on interdisciplinary evidence, she develops a hierarchical generative framework based on principles of probabilistic predictive processing and offers a unifying account of how abstract event-predictive encodings are formed. She further details how the mind both monitors the unfolding events and updates the encoding of such events in memory, focusing on consistency of its predictions. This processing enables the detection of event boundaries necessary for event segmentation and the formation of compact hierarchical generative structures that can be stored in memory. While Kuperberg primarily focuses on event perception and inferring intentions and goals upon observing the actions of other agents, she also offers extensions of her framework to goal-directed behavior and the process of goal-setting from the point of view of the acting agent. Finally, the paper addresses the social implications of successfully formed event-predictive encodings for the perception of novel events and unfamiliar agents. Overall, the paper highlights many important interactions between learning, memory, and the perception and pursuance of environmental interactions, all of which continuously unfold in and are mediated by a developing hierarchical generative event-predictive model, that is, our brain.

5. Conclusion

The contributions in this volume offer a highly interdisciplinary perspective on event-predictive cognition. We suggest that such a perspective is useful—indeed critical—for highlighting and comprehending the significance of event-predictive mechanisms in structuring human cognition. There is a convergence on models of event representation from across many disciplines and theoretical perspectives. The available evidence suggests that well-structured event-predictive encodings and processes critically focus the cognitive mechanisms that deliver the human ability to interact with the environment in a highly versatile, social, and anticipatory goal-directed manner. That is to say, event-predictive encodings and processes are of critical importance in delivering intelligent human behavior.

The implications of this proposal go beyond the theoretical understanding of how the human mind works. Event processing is a crucial component of learning. Our ability to process, categorize, and remember new events depends on the properties of the sensorimotor stream that surrounds those events. Currently irrelevant sensory signals, such as, for example, notifications and messages we receive on our phones while being in and focusing on different events, interfere with the processes that help to consolidate new events in memory. Thus, educational strategies should take the findings related to event segmentation and recall into account. Moreover, even for adults excessive event interruptions are not only disturbing but also cognitively strenuous, calling for effective strategies for dealing with or simply ignoring current non-urgent stimuli.

With respect to artificial intelligence, when the task is to develop systems that become truly intelligent, we suggest that their learning mechanisms should be endowed with inductive biases that tend to develop latent, event-predictive encodings. Such encodings tend to yield compact factored, and partially even causal, explanations of the observed sensorimotor dynamics, they enable planning and reasoning on conceptual and compositionally meaningful levels, and they appear to be well-connectible with language encodings and processing mechanisms. As a result, these neurocognitive machine learning systems may be able to uncover innovative problem solutions and recommendations, and may thus outperform current deep learning, classification-oriented machine learning systems by far.

Acknowledgments

Support from a Feodor-Lynen Grant of the Humboldt Foundation for Martin Butz is acknowledged. Martin Butz and Asya Achimova were also supported by the German Research Foundation (DFG; project number 198647426), via the Research Training Group 1808: Ambiguity—Production and Perception.

References

Bacon, P.-L., Harb, J., & Precup, D. (2017). The option-critic architecture. In S. Singh & S. Markovitch (Eds.), *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, AAAI'17* (pp. 1726–1734). San Francisco, CA: AAAI Press.

Bar, M. (2009). Predictions: A universal principle in the operation of the human brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 1181–1182.

Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.

Barto, A. G., & Mahadevan, S. (2003). Recent advances in hierarchical reinforcement learning. *Discrete Event Dynamic Systems*, *13*, 341–379.

Botvinick, M., Niv, Y., & Barto, A. C. (2009). Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. *Cognition*, *113*, 262–280.

Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Sciences*, *16*, 485–488.

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*, 49–57.

Butz, M. V. (2016). Toward a unified sub-symbolic computational theory of cognition. *Frontiers in Psychology*, *7*, https://doi.org/10.3389/fpsyg.2016.00925

Butz, M. V. (2017). Which structures are out there? Learning predictive compositional concepts based on social sensorimotor explorations. In T. K. Metzinger & W. Wiese (Eds.), *Philosophy and predictive processing*. Frankfurt am Main: MIND Group. https://doi.org/10.15502/9783958573093

Butz, M. V. (2008). How and why the brain lays the foundations for a conscious self. *Constructivist Foundations*, *4*, 1–42.

Butz, M. V., Bilkey, D., Humaidan, D., Knott, A., & Otte, S. (2019). Learning, planning, and control in a monolithic neural event inference architecture. *Neural Networks*, *117*, 135–144.

Butz, M. V., & Kutter, E. F. (2017). *How the mind comes into being: Introducing cognitive science from a functional and computational perspective*. Oxford, UK: Oxford University Press.

Butz, M. V., Sigaud, O., & Gérard, P. (2003a). Anticipatory behavior: Exploiting knowledge about the future to improve current behavior. In M. V. Butz, O. Sigaud, & P. Gérard (Eds.), *Anticipatory behavior in adaptive learning systems: Foundations, theories, and systems* (pp. 1–10). Berlin: Springer Verlag.

Butz, M. V., Sigaud, O., & Gérard, P. (2003b). Internal models and anticipations in adaptive learning systems. In M. V. Butz, O. Sigaud, & P. Gérard (Eds.), *Anticipatory behavior in adaptive learning systems: Foundations, theories, and systems* (pp. 86–109). Berlin: Springer Verlag.

Casati, R., & Varzi, A.(2015). Events. InE. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2015 Edition). Available at: https://plato.stanford.edu/archives/win2015/entries/events/

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Friston, K. (2002). Functional integration and inference in the brain. *Progress in Neurobiology*, *68*, 113–143.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*, 1325–1352.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, *13*, 293–301.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2018). Active inference: A process theory. *Neural Computation*, *29*, 1–49.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge, MA: MIT Press.

Griffiths, T. L., Vul, E., & Sanborn, A. N. (2012). Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science*, *21*, 263–268.

Gumbsch, C., Butz, M. V., & Martius, G. (2019). Autonomous identification and goal-directed invocation of event-predictive behavioral primitives. In *IEEE Transactions on Cognitive and Developmental Systems*. https://doi.org/10.1109/TCDS.2019.2925890

Gumbsch, C., Kneissler, J., & Butz, M. V. (2016). Learning behavior-grounded event segmentations. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 1787–1792). Austin, TX: Cognitive Science Society.

Gumbsch, C., Otte, S., & Butz, M. V. (2017). A computational model for the dynamical learning of event taxonomies. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 452–457). London: Cognitive Science Society.

Herbart, J. F. (1825). *Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik und Mathematik. Zweiter, analytischer Teil [Psychology as a science newly grounded on experience, metaphysics, and mathematics. Second part: Analytics]*. Königsberg, Germany: August Wilhem Unzer.

Hoffmann, J. (1993). *Vorhersage und Erkenntnis: Die Funktion von Antizipationen in der menschlichen Verhaltenssteuerung und Wahrnehmung. [Anticipation and cognition: The function of anticipations in human behavioral control and perception.]*. Göttingen, Germany: Hogrefe.

Hohwy, J. (2013). *The predictive mind*. Oxford, UK: Oxford University Press.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–878.

Jackendoff, R. (2002). *Foundations of language. Brain, meaning, grammar, evolution*. Oxford, UK: Oxford University Press.

König, P., & Krüger, N. (2006). Symbols as self-emergent entities in an optimization process of feature extraction and predictions. *Biological Cybernetics*, *94*, 325–334.

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, *20*, 512–534.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, E253. https://doi.org/10.1017/S0140525X16001837

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press.

Maturana, H. R. (1975). The organization of the living: A theory of the living organization. *International Journal of Man-Machine Studies*, *7*, 313–332.

Pezzulo, G., Butz, M. V., Sigaud, O., & Baldassarre, G. (2009). From sensorimotor to higher-level cognitive processes: An introduction to anticipatory behavior systems. In *Anticipatory behavior in adaptive learning systems: From psychological theories to artificial cognitive systems, LNAI 5499* (pp. 1–9). Berlin: Springer-Verlag.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167–201). Berlin: Springer-Verlag.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research Psychologische Forschung*, *68*, 176–188.

Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, *112*, 181–211.

Vendler, Z. (1957). Verbs and times. *Philosophical Review*, *56*(97), 121.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*, 273–293.

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

Papers in this Topic

Baldwin, D. A., & Kosie, J. E. (2021). How does the mind render streaming experience as events? *Topics in Cognitive Science*, *13*, 79–105.

Bilkey, D. K., & Jensen, C. (2021). Neural markers of event boundaries. *Topics in Cognitive Science*, *13*, 128–131.

Butz, M. V., Achimova, A., Bilkey, D., & Knott, A. (2021). Editors' review and introduction: Event-predictive cognition: From sensorimotor via conceptual to language-based structures and processes. *Topics in Cognitive Science*, *13*, 10–24.

Cooper, R. P. (2021). Action production and event perception as routine sequential behaviors. *Topics in Cognitive Science*, *13*, 63–78.

Elsner, B., & Adam, M. (2021). Infants' Goal Prediction for Simple Action-events: The role of experience and agency cues. *Topics in Cognitive Science*, *13*, 45–62.

Hebblewhite, A., Hohwy, J., & Drummond, T. (2021). Events and machine learning. *Topics in Cognitive Science*, *13*, 243–247.

Hohwy, J., Hebblewhite, A., & Drummond, T. (2021). Events, event-prediction and predictive processing. *Topics in Cognitive Science*, *13*, 252–255.

Knott, A., & Takac, M. (2021). Roles for event representations in sensorimotor experience, memory formation, and language processing. *Topics in Cognitive Science*, *13*, 187–205.

Kuperberg, G. R. (2021). Tea with milk? A hierarchical generative framework of sequential event comprehension. *Topics in Cognitive Science*, *13*, 256–298.

McRae, K., Brown, K. S., & Elman, J. L. (2021). Prediction-based learning and processing of event knowledge. *Topics in Cognitive Science*, *13*, 206–233.

Paulin, M. G., & Cahill-Lane, J. (2021). Events in early nervous system evolution. *Topics in Cognitive Science*, *13*, 25–44.

Shin, Y. S., & DuBrow, S. (2021). Structuring memory through inference-based event segmentation. *Topics in Cognitive Science*, *13*, 106–127.

Stawarczyk, D., Bezdek, M. A., & Zacks, J. M. (2021). Event representations and predictive processing: The role of the midline default network core. *Topics in Cognitive Science*, *13*, 164–186.

Storchak, H., Ehlis, A.-C., & Fallgatter, A. (2021). Action monitoring alterations as indicators of predictive deficits in schizophrenia. *Topics in Cognitive Science*, *13*, 164–186.

Ünal, E., Ji, Y., & Papafragou, A. (2021). From event representation to linguistic meaning. *Topics in Cognitive Science*, *13*, 224–242.

Warren, T., & Dresang, H. C. (2021). Event-predictive cognition: Underspecification and interaction with language. *Topics in Cognitive Science*, *13*, 248–251.