

Structured event complexes in the medial prefrontal cortex support counterfactual representations for future planning

Aron K. Barbey, Frank Krueger and Jordan Grafman

Phil. Trans. R. Soc. B 2009 **364**, 1291-1300
doi: 10.1098/rstb.2008.0315

References

[This article cites 60 articles, 13 of which can be accessed free](#)

<http://rstb.royalsocietypublishing.org/content/364/1521/1291.full.html#ref-list-1>

Subject collections

Articles on similar topics can be found in the following collections

[neuroscience](#) (132 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

To subscribe to *Phil. Trans. R. Soc. B* go to: <http://rstb.royalsocietypublishing.org/subscriptions>

Structured event complexes in the medial prefrontal cortex support counterfactual representations for future planning

Aron K. Barbey^{1,2}, Frank Krueger¹ and Jordan Grafman^{1,*}

¹Cognitive Neuroscience Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Building 10, Room 7D43, 10 Center Drive, MSC 1440, Bethesda, MD 20892-1440, USA

²Department of Psychology, Georgetown University, White-Gravenor Hall 306 37th and O Street, NW Washington, DC 20057-1001, USA

We propose that counterfactual representations for reasoning about the past or predicting the future depend on *structured event complexes* (SECs) in the human prefrontal cortex (PFC; ‘What would happen if *X* were performed in the past or enacted in the future?’). We identify three major categories of counterfactual thought (concerning action versus inaction, the self versus other and upward versus downward thinking) and propose that each form of inference recruits SEC representations in distinct regions of the medial PFC. We develop a process model of the regulatory functions these representations serve and draw conclusions about the importance of SECs for explaining the past and predicting the future.

Keywords: structured event complexes; prefrontal cortex; counterfactual reasoning; planning; prediction; event knowledge

1. INTRODUCTION

Remembering the past and predicting the future depend on the ability to shift from perceiving the immediate environment to an alternative, imagined perspective. Mental models of imagined past events or future outcomes not yet at hand support *counterfactual thinking* (‘What would happen if *X* were performed in the past or enacted in the future?’; Kahneman & Miller 1986; Byrne 2002). The capacity for counterfactual thought enables learning from past experience (Byrne 1997), supports planning and predicting future events (Brase & Barbey 2006; Barbey & Sloman 2007), provides the basis for creativity and insight (Sternberg & Gastel 1989; Thomas 1999; Costello & Keane 2000) and gives rise to emotions and social ascriptions (e.g. guilt, regret, blame) that are central for managing and regulating social behaviour (Landman 1987; Miller & Turnbull 1990; Niedenthal *et al.* 1994; Davis *et al.* 1995; Zeelenberg *et al.* 1998). The neural representation of counterfactual inference draws upon neural systems for constructing mental models of the past and future, incorporating prefrontal and medial temporal lobe structures (Tulving & Markowitsch 1998; Fortin *et al.* 2002). In this article, we develop an integrative cognitive neuroscience framework for understanding counterfactual reasoning on the basis of *structured event complexes* (SECs) in the human prefrontal cortex (PFC).

We begin by reviewing the biology and structure of the human PFC and introduce a cognitive neuroscience framework for the representation of event

knowledge within the PFC. We then survey recent neuroscience evidence in support of the SEC framework and establish the role of distinct PFC subregions in the representation of specific forms of event knowledge. After reviewing the cognitive and neural foundations of the SEC framework, we show how this approach accounts for counterfactual reasoning. We identify three major categories of counterfactual inference (concerning action versus inaction, the self versus other and upward versus downward thinking) and review neuroscience evidence for their representation in distinct regions of the medial PFC. Finally, we draw conclusions about the importance of SECs for learning from past experience, planning and predicting future events, creativity and insight, and the management and regulation of social behaviour.

2. NEUROBIOLOGY OF THE HUMAN PREFRONTAL CORTEX

SECs are representations composed of goal-oriented sequences of events involved in executing, planning and monitoring action. We briefly review the biology and structure of the human PFC, providing evidence to support our proposal that the PFC stores cognitive representations intimately concerned with goal-directed action.

The PFC can be divided into ventromedial and dorsolateral regions, each of which is associated with posterior and subcortical brain regions. The ventromedial PFC (vmPFC) has reciprocal connections with brain regions that are associated with emotional processing (amygdala), memory (hippocampus) and higher order sensory processing (temporal visual association areas), as well as with the dorsolateral PFC (dlPFC).

* Author for correspondence (grafmanj@ninds.nih.gov).

One contribution of 18 to a Theme Issue ‘Predictions in the brain: using our past to prepare for the future’.

The dlPFC has reciprocal connections with brain regions that are associated with motor control (basal ganglia, premotor cortex, supplementary motor area), performance monitoring (cingulate cortex) and higher order sensory processing (association areas, parietal cortex). The vmPFC is well suited to support functions involving the integration of information about emotion, memory and environmental stimuli, and the dlPFC to support the regulation of behaviour and control of responses to environmental stimuli.

PFC neurons are particularly able to fire over extended periods of time (Levy & Goldman-Rakic 2000) and across events (Fuster & Alexander 1971; Bodner *et al.* 1996). This indicates that the PFC can maintain stimulus representations across time, enabling a subject to engage in behaviour to achieve long-term goals. In addition, pyramidal cells in the macaque PFC are more spinous—and therefore can handle more excitatory inputs—than other cortical pyramidal cells (Elston 2000). This is one structural explanation for the PFC's ability to integrate inputs from many sources and to implement complex behaviours. The monkey's PFC contains cells that respond to both internally generated and observed behaviours—these have been termed *mirror neurons* (Gallese *et al.* 1996). Similar regions have been shown to be activated in humans when observing and performing actions (Grafton *et al.* 1996; Rizzolatti 1996). These data support a role for the PFC in the representation of action. Furthermore, Williams and colleagues have suggested that abnormal development of the PFC might lead to impaired social behaviour (Williams *et al.* 2001), which can also be caused by PFC damage later in life.

It is thought that the dlPFC evolved from motor regions and developed much later than the vmPFC (Fuster 1997). Motor regions store motor programmes; therefore, it seems reasonable that the functions of the 'newer' PFC regions would be related to those of the older ones, providing a representational basis for goal-directed action.

In summary, the connectivity of PFC regions, physiological properties of its neurons and evolutionary principles are strongly suggestive of its role in the integration of sensory and memory information and in the representation and control of actions and behaviour. Along with the extended firing of neurons, specialized neural systems were developed that enabled the parsing and encoding of these behaviours into sequentially linked, but individually recognizable, events. At the broadest level, events are parsed into subcomponents consisting of an activity that signals the onset of the event, followed by a series of activities performed to achieve the desired goal, and a final activity resulting in event completion. Events are further characterized by their semantic content, temporal duration and the number of component activities they entail (Zacks & Tversky 2001; Zacks *et al.* 2001).

We propose that the structure of event knowledge can be conceptualized as a 'representation' or a unique form of knowledge that, when activated, corresponds to a dynamic brain state signified by the strength and pattern of neural activity in a local brain region. In this

sense, over the course of evolution, the PFC became capable of representing knowledge of more complex behaviours. We have labelled these representational units within the PFC *structured event complexes* (SECs; Barbey *et al.* *in press a, b*).

3. STRUCTURED EVENT COMPLEX THEORY

An SEC is a goal-oriented set of events that is structured in sequence and represents event features (including agents, objects, actions, mental states and background settings), social norms of behaviour, ethical and moral rules and temporal event boundaries (see also Barbey *et al.* *in press a, b*). Aspects of SECs are represented independently, but are encoded and retrieved as an episode. The SEC theory is a representational framework that motivates specific predictions regarding the properties and localization of SECs within the PFC (figure 1). We review principal elements of the SEC theory before turning to an assessment of its neurobiological predictions.

(a) *Neural architecture*

SECs are encoded and activated on the basis of simulation mechanisms (Damasio 1989; Barsalou *et al.* 2003a,b). A large body of neuroscience evidence demonstrates that experience in the physical and social world activates feature detectors in the relevant feature maps of the brain. During visual processing of a face, for example, some neurons fire for edges and planar surfaces, whereas others fire for colour, configural properties and movement. The global pattern of activation across this hierarchically organized distributed system represents the entity in vision (Zeki 1993; Palmer 1999). Analogous patterns of activation on other sensory modalities represent how the face might sound and feel. Activation in the motor system similarly represents responses to the face, such as the formation of a facial expression and approach/avoidance behaviour. A similar mechanism underlies the introspective states that arise while interacting with an entity. For example, activation patterns in the amygdala and orbitofrontal areas represent emotional reactions to social stimuli. Much neuroscience evidence documents the structure of feature maps across modalities and the states that arise in them.

When a pattern becomes active in a feature map during perception or action, conjunctive neurons in an association area capture the pattern for later cognitive use. For example, conjunctive neurons in the visual system capture the pattern active for a particular face. A population of conjunctive neurons together codes a particular pattern, with each individual neuron participating in the coding of many different patterns. Damasio (1989) called these association areas *convergence zones* and proposed that they exist at multiple hierarchical levels in the brain, ranging from posterior to anterior. Most locally, convergence zones near a modality capture activation patterns within it. Association areas near the visual system, for example, capture patterns there, whereas association areas near the motor system capture patterns in this local region. Downstream in more anterior regions higher association areas—including temporal and frontal regions—integrate activation across modalities.

structured event complex prefrontal cortex	
hemispheric lateralization	
left PFC <i>single-event integration</i> meaning and features between single adjacent events to code for event boundaries	right PFC <i>multiple-event integration</i> meaning and features across events to obtain goal of sequence
predictability	
lateral PFC <i>adaptive SECs</i> frequently modified sequences that are used to adapt to special circumstances	medial PFC <i>predictable SECs</i> overlearned sequences that have clearly definable goals and all behavioural rules available
complexity	
anterior PFC <i>complex SECs</i> detailed information about event sequences	posterior PFC <i>simple SECs</i> sparse information about event sequences
category specificity	
dorsolateral PFC <i>planning and action</i> event sequences representing mechanistic plans and actions	ventromedial PFC <i>behavioural norms and scripts</i> event sequences representing social norms and scripts

Figure 1. SEC framework. The representational forms of structured event complexes (SEC) and their proposed localizations within the prefrontal cortex (PFC).

According to the SEC framework, event knowledge is represented by higher order convergence zones localized within particular regions of the PFC (figure 1). Once a set of conjunctive neurons within the PFC captures feature maps (representing components of event knowledge, social norms, ethical and moral rules and temporal event boundaries), the set can later activate the pattern in the absence of bottom-up stimulation. For example, on entering a familiar situation and recognizing it, an SEC that represents the situation becomes active. Typically not all of the situation is perceived initially. A relevant person, setting or event may be perceived, which then suggests that a particular situation is about to play out. It is in the agent's interests to anticipate what will happen next so that optimal actions can be executed. The agent must draw inferences that go beyond the information given (Griffin & Ross 1991). The SEC that becomes active constitutes a rich source of social inference supporting the planning, execution and monitoring of action. The SEC can be viewed as a distributed pattern representing components of event knowledge (i.e. as a complex configuration of multimodal components that

represent the situation). Because part of this pattern matched the current situation initially, the larger pattern became active in memory. The remaining parts of the pattern—not yet observed in the situation—constitute inferences, namely educated guesses about what might occur next. Because the remaining parts co-occurred frequently with the perceived parts in previous situations, inferring the remaining parts from the perceived parts is reasonable. As a partially viewed situation activates an SEC, the SEC completes the pattern that the situation suggests.

To the extent that the SEC is entrenched in memory, pattern completion is likely to occur at least somewhat automatically. As a situation is experienced repeatedly, its simulated components and the associations linking them increase in potency. Thus, when one component is perceived initially, these strong associations complete the pattern automatically. Consider the example of meeting with a colleague. Her face, clothing and bodily mannerisms initially match modality-specific simulations in one or more SECs that have become entrenched in memory. Once one of these wins the activation process, it provides inferences via pattern

completion, such as actions that the colleague is likely to take, actions that the perceiver typically takes, affective states that are likely to result and so forth. The unfolding of such inferences—realized as an SEC—produces social prediction (for a cognitive neuroscience review of simulation mechanisms in reasoning, see Barbey & Barsalou 2009; Barbey *et al.* in press *a, b*; Patterson & Barbey in press).

(b) *Sequence structure*

SECs integrate modality-specific components of event knowledge, providing the semantic and temporal structure underlying goal-directed action. At the broadest level, SECs link event subcomponents consisting of an activity that signals the onset of the event (e.g. ‘hearing the morning alarm clock’), followed by a series of activities performed to achieve the desired goal (e.g. ‘waking up’, ‘getting out of bed’), and a final activity resulting in event completion (e.g. ‘arriving at work’). The temporal structure of SECs further obeys cultural and individual constraints, reflecting socio-cultural norms of appropriate behaviour (e.g. in the United States, people typically shower in the morning daily) and personal preferences concerning the temporal order and frequency of performed activities (e.g. the individual preference to shower in the morning and at night daily). The semantic and temporal structure of event knowledge supports goal-directed action in dynamic environments, enabling the on-line modification of specific activities (e.g. owing to changing circumstances) and the simulation of only those event components necessary for goal achievement in the present context (e.g. beginning at various stages in the event sequence, returning to earlier stages, skipping unnecessary activities owing to time pressure).

(c) *Goal orientation*

The semantic and temporal structure of SECs derives from event goals that provide the basis for the selection, temporal ordering and execution of activities underlying an event. Some SECs are *well structured*, with clearly defined goals and cognitive and behavioural action sequences that are available for goal achievement. For example, individuals with a well-structured SEC for ‘eating in a restaurant’ are quite confident that once they have been seated at a table and have read the menu, someone will appear to take their order.

By contrast, some SECs are *ill-structured*, requiring the individual to adapt to unpredictable circumstances by constructing novel or *ad hoc* goals, and selecting appropriate action sequences on-line (Barsalou 1991). For example, if someone sees that a person entering a bank is wearing a ski mask and carrying a gun, one can make sense of these events by completing the activated ‘bank robbery’ SEC to access further components of event knowledge (concerning relevant agents, objects, actions, mental states and background settings).

(d) *Binding*

Multiple SECs are activated to support the events of our daily life and therefore it is likely that these representations (such as events within an SEC) can be activated in sequence or, additionally, in a cascading

or parallel manner. Event components interact and give rise to SECs through at least three binding mechanisms: sequential binding, proposed for linking multiple SECs within the PFC (Weingartner *et al.* 1983); temporal binding among anatomically integrated regions representing event subcomponents in the posterior cortex (Engel & Singer 2001); and third-party binding of anatomically loosely connected regions whose activity is synchronized via the hippocampus (Weingartner *et al.* 1983; O’Reilly & Rudy 2000).

(e) *Hierarchical structure*

Given the slow development of the PFC during childhood, individual events are probably initially represented as independent memory units. For example, SECs associated with ‘kitchen’ and ‘school cafeteria’ cluster around the event ‘eat meal’, whereas ‘car’ and ‘school bus’ cluster around the event ‘travel to new location’. Later in development, these primitive SECs expand into large multi-event units, based on repeated exposure and goal-directed action. In addition, the boundaries of event sequences become more firmly established, leading to a well-structured SEC. Thus, in adulthood, SECs will range from specific episodes to more abstract SECs that can be applied to a variety of situations (Barsalou & Wiemer-Hastings 2005). For example, the domain ‘eat meal’ includes specific episodes representing evenings at a particular restaurant, SECs representing the actions and themes of how to behave at different types of restaurants, in addition to more abstract SECs representing actions and themes related to ‘eating’ that apply to a broad range of situations (e.g. at ‘restaurants’, ‘parties’, ‘picnics’, ‘baseball games’).

4. COUNTERFACTUAL THOUGHT

We propose that SECs provide the basis for counterfactual reasoning about past and future events and develop a process model of the regulatory functions these representations serve. Our analysis reveals three major categories of counterfactual inference and grounds each form of inference in SEC representations within distinct regions of the mPFC.

Counterfactual thinking involves mentally undoing the present state of affairs and imagining alternative realities ‘if only’ different decisions were made or actions taken (Kahneman & Miller 1986; Byrne 2002). We propose that counterfactual thought depends on mental models of alternative possibilities represented in the form of SECs. For example, the counterfactual inference that ‘If we chose to sail the Mediterranean rather than continue writing, then you would not be reading this article now’ draws upon SEC knowledge, including the representation of relevant agents (e.g. the authors), objects (e.g. a sailing boat), actions (e.g. sailing), mental states (e.g. freedom) and background settings (e.g. the Mediterranean Sea). Simulations of the representational elements of SEC knowledge provide the basis for evaluating the consequences of alternative courses of action, with the simulation of the authors ‘sailing the Mediterranean’ resulting in the failure to complete this article.

A growing body of research demonstrates the importance of counterfactual inference for generating predictions about the future, supporting representations of unknown future possibilities critical for planning and decision making (e.g. 'How well would the Cubs perform next season if the manager would have acquired key players in the off season?'; Brase & Barbey 2006; Barbey & Sloman 2007). Predictions about the future are supported by modifying a factual event (e.g. 'If the manager acquired key players in the off season...') and considering likely future consequences (e.g. '... would the team perform well next year?').

The SEC framework advocates a theory of *motivated thinking* (Dunning 1999; De Dreu *et al.* 2006), proposing that *drives, needs, desires, motives* and *goals* structure and organize components of event knowledge and profoundly influence judgement and decision making. According to this framework, the primary role of counterfactual thought is to support emotions and social ascriptions that are central for managing and regulating social behaviour. In particular, counterfactual reasoning enables the representation of guilt, regret and blame, which are central for adaptive social behaviour (Landman 1987; Miller & Turnbull 1990; Niedenthal *et al.* 1994; Davis *et al.* 1995; Zeelenberg *et al.* 1998). For example, the counterfactual inference that 'The university would offer her a higher salary (in the past or future) if she were a man' gives rise to feelings of guilt and regret (for the observed gender inequity) that promote behavioural change, and enable the assessment of blame (held by university policymakers) to support planning and decision making (e.g. to apply for positions at other universities). Counterfactual inference therefore enables an assessment of the consequences of alternative decisions or action sequences central for the representation of guilt, regret and blame.

The neural representation of emotions and social ascriptions is distributed throughout the mPFC and is integrated with posterior knowledge networks via binding mechanisms in the medial temporal lobe (Tulving & Markowitsch 1998; Fortin *et al.* 2002; Moll & de Oliveira-Souza 2007). This distributed pattern of multimodal information (e.g. representing agents, objects, actions, mental states, background settings) gives rise to mental models for counterfactual inference.

(a) *Process model*

According to the SEC framework, counterfactual thought is deeply connected to drives, needs, desires, motives and goals, and provides the basis for regulatory mechanisms that keep behaviour on track, particularly within social interactions. We propose that counterfactual thought depends on SEC representations and operates according to the following interactive process model (figure 2) as follows.

- (i) Counterfactual thoughts are activated when a problem is encountered or anticipated in the future. Failure to achieve the desired goal or the anticipation of goal failure in the future typically initiates counterfactual thinking (e.g. owing to negative emotions, the desire for rewards associated with goal achievement).

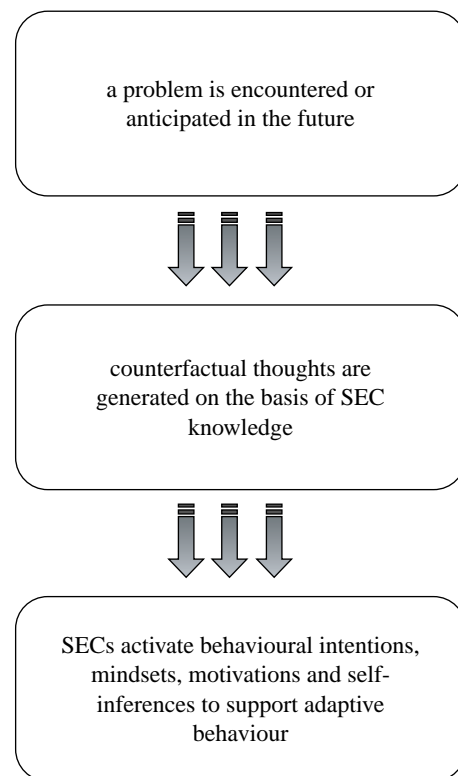


Figure 2. A process model of the regulatory functions underlying counterfactual thought.

- (ii) Counterfactual thoughts are generated from causal implications represented by SECs in the form of events (agents, objects, actions, mental states and background settings) that lead to a desired goal state (for a review of psychological theories of causal representation and reasoning, see Barbey & Wolff 2006, 2007, submitted; Chaigneau & Barbey 2008; Patterson & Barbey *in press*; Sloman *et al.* *in press*).
- (iii) SECs activate corresponding behavioural intentions (e.g. to perform a particular action), mindsets (e.g. to focus on a particular class of events), motivations (e.g. to modulate one's desire for a particular outcome) and/or self-inferences (e.g. to monitor one's public image) which initiate corrective behaviour.

To the extent that such behaviour alleviates the original problem, this mechanism is effective in regulating behaviour in terms of goal pursuit (for a review of medical health applications, see Gilbar & Heyroni 2007; Wrosch *et al.* 2007).

(b) *Categories of counterfactual inference*

The proposed role of SECs in counterfactual thought motivates the prediction that this form of inference will depend on core elements of SECs, which fundamentally represent *actions* performed by *agents* leading to an observed *outcome*. The psychological literature on counterfactual thought supports this prediction, identifying three major categories of counterfactual inference corresponding to core components of SEC representations.

(c) Action knowledge

One broad distinction represents counterfactual thought about *action* versus *inaction*, or the addition versus subtraction of an action from the present state (Roese *et al.* 1999). For example, the counterfactual inference that ‘She should never go out the night before an exam’ represents the addition of an action (‘going out’), whereas the inference that ‘He should always read the instructions carefully’ represents the removal of an action (‘reading carefully’). This form of counterfactual thought is central for evaluating consequences of carrying out or failing to perform specific actions (in the past or future).

(d) Agent knowledge

A second major category of counterfactual inference represents reasoning about the *self* versus *other* (Mandel 2003). For example, the counterfactual inference that ‘Problems would be avoided if I attended the meeting’ represents features of the self, whereas the inference that ‘Your skills would improve if you played more often’ embodies features of others. Counterfactual reasoning about the self versus other provides the basis for adaptive social behaviour and inferring the connection between specific mental states and particular patterns of behaviour (e.g. ‘She would not have left early if she had wanted to talk with you’).

(e) Outcome knowledge

A third major category represents the comparison of a current outcome with a *better or worse alternative* (Roese & Olson 1995). For example, the counterfactual inference that ‘She should accept the job with the higher salary’ represents an upward inference about a better alternative, whereas the observation that ‘Other people with her qualifications earn much less than she does’ represents a downward inference about a worse alternative. Counterfactual reasoning about better versus worse outcomes is critical for learning from the past and assessing alternative courses of action in the future.

The reviewed categories of counterfactual inference embody core features of SEC knowledge, enabling adaptive social behaviour on the basis of actions, agents and event outcomes.

5. NEUROSCIENCE REVIEW

We review neuroscience findings in support of the SEC framework, providing evidence to confirm the representational role of the human PFC and to support the role of SECs in counterfactual thought. The representational aspects of SECs and their proposed localizations within the PFC are summarized in [figure 1](#) (for a review of further evidence in support of the SEC framework, see Barbey *et al.* *in press a, b*; Krueger *et al.* *in press*).

(a) Category specificity of the PFC

The subdivision of the PFC into neuroanatomically distinct regions designed to process specific forms of knowledge supports the proposal that SEC representations are stored within particular regions of the PFC on a content-specific basis ([figure 1](#)). Converging

evidence is provided by lesion studies demonstrating selective impairments for social and reward-related behaviour in vmPFC lesion patients (Dimitrov *et al.* 1999; Milne & Grafman 2001), and impairments for mechanistic planning in dlPFC patients (Burgess *et al.* 2000; Goel & Grafman 2000).

Our research group conducted a PET study providing further evidence to support the representation of domain-specific SECs for *non-emotional* versus *emotional* event knowledge within the PFC (Partiot *et al.* 1995). The employed non-emotional task asked subjects to ‘imagine silently the sequence of events and feelings concerned with preparation and dressing before (their) mother comes over for dinner’. By contrast, subjects in the emotional task were asked to ‘imagine silently the sequence of events and feelings concerned with preparation and dressing to go to (their) mother’s funeral’. Consistent with the domain-specific predictions of the SEC framework, distinct patterns of neural activity were observed when subjects assessed non-emotional versus emotional scripts. Non-emotional scripts activated the right superior frontal gyrus (BA 8), the bilateral middle gyrus (BA 8 and 9) and the medial frontal gyrus (BA 6 and 10), whereas emotional scripts recruited the left anterior cingulate (BA 24 and 32), bilateral medial frontal gyrus (BA 8 and 9) and anterior medial temporal lobe (BA 21).

Employing fMRI, we further demonstrated that *social* versus *non-social* SECs depend on a distinct representational topography within the PFC (Wood *et al.* 2003). We applied a modified go/no-go paradigm in which subjects classified individual words (e.g. ‘menu’, ‘order’) or phrases (e.g. ‘read the menu’, ‘order the food’) according to one of two focal categories (social versus non-social). Social activities recruited the left superior frontal gyrus (BA 8 and 9), whereas non-social activities engaged the right superior frontal gyrus (BA 8), the left medial frontal gyrus (BA 6) and the bilateral anterior cingulate (BA 25). Despite the large body of evidence to support the role of the orbitofrontal cortex (OFC) in social processing (Fuster 1997; Milne & Grafman 2001), activation in this region was not observed. Further inspection of the functional images demonstrated signal dropout in the OFC, limiting conclusions drawn concerning the role of this region in the storage of social SECs.

To further investigate this issue, we conducted a lesion study in which patients with PFC lesions and matched controls performed the classification task of Wood *et al.* (2003, 2005a). Subjects classified individual words (e.g. menu, order) or phrases (e.g. read the menu, order the food) as representing *social* versus *non-social* events. Patients with damage to the right OFC demonstrated cognitive impairments in the accessibility of script and semantic representations of social (rather than non-social) activities, providing evidence to support the role of the OFC in social processes.

In a subsequent fMRI study, we applied multidimensional scaling to assess the psychological structure of event knowledge and its neural representation within particular regions of the PFC (Wood *et al.* 2005b). Multidimensional scaling revealed three psychological dimensions underlying event knowledge

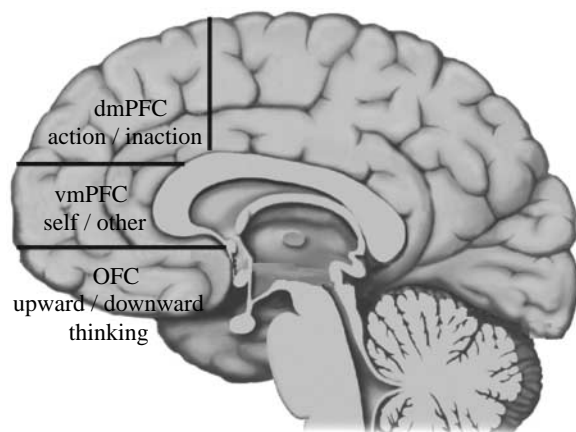


Figure 3. Neural predictions of the SEC theory of counterfactual thought. The dorsomedial prefrontal cortex (dmPFC) represents counterfactual reasoning about *action* versus *inaction*, the ventromedial prefrontal cortex (vmPFC) represents counterfactual thoughts directed towards the *self* versus *other* and the orbitofrontal cortex (OFC) represents *upward* versus *downward* counterfactual thinking.

(*engagement, social valence and experience*). To investigate the neural correlates of the identified psychological dimensions, we conducted an fMRI experiment in which subjects classified each event according to whether or not it represented a social activity. Parametric analyses of event-related fMRI data were applied to investigate brain regions whose activity was modulated by the three psychological components of event knowledge. The results demonstrated that the psychological structure of event knowledge is broadly organized along dimensions that are represented within distinct regions of the human PFC, with the *experience* dimension recruiting the medial PFC (BA 10), the *engagement* dimension activating the left OFC (BA 47) and the *social valence* dimension engaging the amygdala and the right OFC (BA 11 and 47).

In summary, the reviewed studies provide evidence to support our proposal that category-specific SECs are stored within distinct regions of the PFC (see also Barbey *et al.* in press *a, b*; Krueger *et al.* in press).

(b) Counterfactual thought

We propose that counterfactual reasoning is characterized by three major forms of inference that each recruit distinct regions of the mPFC (figure 3). According to this framework, counterfactual thinking depends on category-specific SECs within the mPFC, which provide key representational elements within a larger network of anatomically connected prefrontal and posterior regions supporting counterfactual thought.

(c) Action versus inaction

According to the SEC framework, counterfactual reasoning about action versus inaction preferentially recruits the dorsomedial PFC (dmPFC). Several neuroscience studies have implicated the dmPFC in the continuous internal monitoring of action (Botvinick *et al.* 2004). Barch *et al.* (2001) report an extensive meta-analysis of functional imaging studies that included data from a broad range of action monitoring tasks (e.g. involving the inhibition of

prepotent responses) that recruited the dmPFC. Along the same lines, Walton *et al.* (2004) observed activity in the dmPFC when participants monitored the outcome of self-selected actions. These findings suggest that the dmPFC is critical for monitoring the addition versus subtraction of actions for counterfactual reasoning.

(d) Self versus other

We propose that counterfactual thought involving the self versus others recruits the vmPFC. A large body of neuroscience evidence supports this proposal, demonstrating that the vmPFC represents components of self-knowledge, person knowledge and mentalizing. Beldarrain *et al.* (2005) demonstrated impairments in self-generated counterfactual thought in vmPFC lesion patients. Converging evidence is provided by Macrae *et al.* (2004), who observed activation in the vmPFC when participants evaluated the self-relevance of specific personality traits. Similarly, Ochsner *et al.* (2004) found activation in the vmPFC when participants monitored their emotional states.

Recruitment of the vmPFC is also observed in studies that assess person knowledge more broadly (applying to others as well as the self). Mitchell *et al.* (2002) reported activation in this region when participants judged whether a presented adjective applied to a person (rather than an inanimate object). Consistent with these findings, Schmitz *et al.* (2004) observed activation in the vmPFC when participants thought about themselves or a close friend.

Finally, extensive neuroscience evidence implicates the vmPFC in the process of representing another person's psychological perspective (i.e. 'mentalizing'). For example, Fletcher *et al.* (1995) and Goel *et al.* (1995) reported activation in the vmPFC when participants read social scripts in which the psychological perspective of fictional characters was inferred.

(e) Upward versus downward thinking

We propose that counterfactual reasoning about upward (better) versus downward (worse) outcomes recruits the OFC, which is widely implicated in the processing of event outcomes associated with rewards or penalties. Elliott *et al.* (2000) propose that the OFC is involved in monitoring reward and serves to guide behaviour in terms of the value of possible outcomes. Walton *et al.* (2004) found that the activity in the OFC was elicited by the need to monitor the outcomes of externally guided actions. Similarly, Coricelli *et al.* (2005) found that activity in the OFC correlated with the amount of anticipated regret associated with a decision. In summary, the reviewed findings suggest that the OFC provides the basis for counterfactual reasoning about upward (better) versus downward (worse) outcomes.

6. CONCLUSION

We have introduced a 'representational' theory of PFC function in accord with the structure, neurophysiology and connectivity of the PFC, and the modern cognitive neuroscience view that elements of knowledge are represented within functionally localized brain regions. The reviewed evidence in support

of the SEC framework confirms the importance and uniqueness of the human PFC for representing knowledge in the form of cognitive events and action sequences.

We have further advocated for the representational basis of SECs in counterfactual thought, reviewing evidence to support the role of specific regions of the mPFC in the representation of particular forms of counterfactual inference. According to this framework, SECs in the mPFC represent components of event knowledge (agents, objects, actions, mental states and background settings) that are essential for constructing mental models of past or future events and assessing the consequences of alternative courses of action. Our findings underscore the importance of SECs for high-level cognition more broadly, supporting their role in the construction of mental models and the simulation of alternative possibilities for learning from past experience (e.g. 'We should not have gone to Las Vegas'; Byrne 1997), planning and predicting future events (e.g. 'Next time we should go sailing'; Brase & Barbey 2006; Barbey & Sloman 2007), creativity and insight (e.g. 'Perhaps we could write while sailing'; Sternberg & Gastel 1989; Thomas 1999; Costello & Keane 2000) and adaptive social behaviour (e.g. supported by regulatory mechanisms based on representations of guilt, regret and blame; Landman 1987; Miller & Turnbull 1990; Niedenthal *et al.* 1994; Davis *et al.* 1995; Zeelenberg *et al.* 1998).

In conclusion, we believe SECs are the key to understanding the human ability to represent mental models of events, which guide the selection of goal-directed action sequences and the on-line updating of behaviour based on past outcomes or anticipated future events. When stored as memories, SECs provide a link between past, current and future activities, enabling explanatory and predictive inferences that give rise to adaptive behaviour and issue significant advantages for our species. Our review demonstrates that there is now substantial evidence to suggest that studying the nature of SEC representations is a competitive and promising way to characterize the components of event knowledge stored within the human PFC. Although SECs must coordinate with representations and processes stored in other brain regions—requiring hippocampal and related structure-binding processes for the sense of an episode to emerge in consciousness—the elusive scientific characterization of knowledge stored within the PFC remains the key missing part of the puzzle. We believe that the evidence collected so far has brought us one step closer to such an understanding of the contribution of the PFC to future planning.

The authors are supported by the NINDS Intramural Research Program.

REFERENCES

Barbey, A. K. & Barsalou, L. W. 2009 Reasoning and problem solving: models. In *New encyclopedia of neuroscience*, vol. 3 (eds L. Squire, T. Albright, F. Bloom, F. Gage & N. Spitzer), pp. 35–43. Oxford, UK: Elsevier.

- Barbey, A. K. & Sloman, S. A. 2007 Base-rate respect: from ecological rationality to dual processes. *Behav. Brain Sci.* **30**, 241–297. (doi:10.1017/S0140525x07001653)
- Barbey, A. K. & Wolff, P. 2006 Causal reasoning from forces. In *Proc. 28th Annual Conf. Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.
- Barbey, A. K. & Wolff, P. 2007 Learning causal structure from reasoning. In *Proc. 29th Annual Conf. Cognitive Science Society*. Hillsdale, NJ: Erlbaum Press.
- Barbey, A. K. & Wolff, P. Submitted. Composing causal relations in force dynamics.
- Barbey, A. K., Krueger, F. & Grafman, J. In press *a* An evolutionarily adaptive neural architecture for social exchange. *Trends Neurosci.*
- Barbey, A. K., Krueger, F. & Grafman, J. In press *b* Structured event complexes in the prefrontal cortex mediate inductive and deductive inference. In *Encyclopedia of cognitive science*, New York, NY: Wiley Press.
- Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J. & Snyder, A. 2001 Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cereb. Cortex* **11**, 837–848. (doi:10.1093/cercor/11.9.837)
- Barsalou, L. W. 1991 Deriving categories to achieve goals. In *The psychology of learning and motivation: advances in research and theory* (ed. G. H. Bower), pp. 1–64. San Diego, CA: Academic Press.
- Barsalou, L. W. & Wiemer-Hastings, K. 2005 Situating abstract concepts. In *Grounding cognition: the role of perception and action in memory, language, and thought* (eds D. Pecher & R. Zwaan), pp. 129–163. New York, NY: Cambridge University Press.
- Barsalou, L. W., Simmons, W. K., Barbey, A. K. & Wilson, C. D. 2003*a* Grounding conceptual knowledge in modality-specific systems. *Trends Cogn. Sci.* **7**, 84–91. (doi:10.1016/S1364-6613(02)00029-3)
- Barsalou, L. W., Niedenthal, P. M., Barbey, A. K. & Ruppert, J. 2003*b* Social embodiment. In *The psychology of learning and motivation* (ed. B. Ross), pp. 43–92. San Diego, DA: Academic Press.
- Beldarrain, M., Garcia-Monco, J., Astigarraga, E., Gonzalez, A. & Grafman, J. 2005 Only spontaneous counterfactual thinking is impaired in patients with prefrontal cortex lesions. *Cogn. Brain Res.* **24**, 723–726. (doi:10.1016/j.cogbrainres.2005.03.013)
- Bodner, M., Kroger, J. & Fuster, J. M. 1996 Auditory memory cells in dorsolateral prefrontal cortex. *Neuroreport* **7**, 1905–1908. (doi:10.1097/00001756-199608120-00006)
- Botvinick, M. M., Cohen, J. D. & Carter, C. S. 2004 Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* **8**, 539–546. (doi:10.1016/j.tics.2004.10.003)
- Brase, G. & Barbey, A. K. 2006 Mental representations of statistical information. In *Advances in psychological research* (ed. A. Columbus), pp. 91–113. New York, NY: Nova Science Publishers.
- Burgess, P. W., Veitch, E., de Lacy Costello, A. & Shallice, T. 2000 The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia* **38**, 848–863. (doi:10.1016/S0028-3932(99)00134-7)
- Byrne, R. M. J. 1997 Cognitive processes in counterfactual thinking about what might have been. In *The psychology of learning and motivation, advances in research and theory* (ed. D. Medin), pp. 105–154. San Diego, DA: Academic Press.
- Byrne, R. M. J. 2002 Mental models and counterfactual thinking. *Trends Cogn. Sci.* **6**, 405–445. (doi:10.1016/S1364-6613(02)01974-5)
- Chaigneau, S. & Barbey, A. K. 2008. Assessing psychological theories of causal meaning and inference. In *Proc. Thirtieth Annual Science Society*. Hillsdale, NJ: Erlbaum.

- Coricelli, G., Critchley, H. D., Joffily, M., O'Doherty, J. P., Sirigu, A. & Dolan, R. J. 2005 Regret and its avoidance: a neuroimaging study of choice behavior. *Nat. Neurosci.* **8**, 1255–1262. (doi:10.1038/nn1514)
- Costello, F. J. & Keane, M. T. 2000 Efficient creativity: constraint guided conceptual combination. *Cogn. Sci.* **24**, 299–349. (doi:10.1016/S0364-0213(00)00020-3)
- Damasio, A. R. 1989 Time-locked multiregional retro-activation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition* **33**, 25–62. (doi:10.1016/0010-0277(89)90005-X)
- Davis, C. J., Lehman, D. R., Wortman, C. B., Silver, R. C. & Thompson, S. C. 1995 The undoing of traumatic life events. *Pers. Soc. Psychol. Bull.* **21**, 109–124. (doi:10.1177/0146167295212002)
- De Dreu, C. K. W., Beersma, B., Stroebe, K. & Euwema, M. C. 2006 Motivated information processing, strategic choice, and the quality of negotiated agreement. *J. Pers. Soc. Psychol.* **6**, 927–943. (doi:10.1037/0022-3514.90.6.927)
- Dimitrov, M., Phipps, M., Zahn, T. & Grafman, J. 1999 A thoroughly modern gage. *Neurocase* **5**, 345–354. (doi:10.1080/13554799908411987)
- Dunning, D. 1999 A newer look: motivated social cognition and the schematic representation of social concepts. *Psychol. Inq.* **10**, 1–11. (doi:10.1207/s15327965pli1001_1)
- Elliott, R., Dolan, R. J. & Frith, C. D. 2000 Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cereb. Cortex* **10**, 308–317. (doi:10.1093/cercor/10.3.308)
- Elston, G. N. 2000 Pyramidal cells of the frontal lobe: all the more spinous to think with. *J. Neurosci.* **20**, RC95.
- Engel, A. K. & Singer, W. 2001 Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* **5**, 16–25. (doi:10.1016/S1364-6613(00)01568-0)
- Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J. & Frith, C. D. 1995 Other minds in the brain: a functional imaging study of 'theory of mind' in story comprehension. *Cognition* **57**, 109–128. (doi:10.1016/0010-0277(95)00692-R)
- Fortin, N. J., Agster, K. & Eichenbaum, H. B. 2002 Critical role of the hippocampus in memory for sequences of events. *Nat. Neurosci.* **5**, 458–562. (doi:10.1038/nn834)
- Fuster, J. M. 1997 *The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe*. New York, NY: Raven.
- Fuster, J. M. & Alexander, G. E. 1971 Neuron activity related to short-term memory. *Science* **173**, 652–654. (doi:10.1126/science.173.3997.652)
- Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. 1996 Action recognition in the premotor cortex. *Brain* **119**, 593–609. (doi:10.1093/brain/119.2.593)
- Gilbar, O. & Hevroni, A. 2007 Counterfactuals, coping strategies and psychological distress among breast cancer patients. *Anxiety Stress Coping* **20**, 382–392. (doi:10.1080/10615800701384439)
- Goel, V. & Grafman, J. 2000 The role of the right prefrontal cortex in ill-structured problem solving. *Cogn. Neuropsychol.* **17**, 415–436. (doi:10.1080/026432900410775)
- Goel, V., Grafman, J., Sadato, N. & Hallett, M. 1995 Modeling other minds. *Neuroreport* **6**, 1741–1746. (doi:10.1097/00001756-199509000-00009)
- Grafton, S. T., Arbib, M. A., Fadiga, L. & Rizzolatti, G. 1996 Localization of grasp representations in humans by PET. II. Observation compared with imagination. *Exp. Brain Res.* **112**, 103–111. (doi:10.1007/BF00227183)
- Griffin, D. W. & Ross, L. 1991 Subjective construal, social inference, and human misunderstanding. *Adv. Exp. Soc. Psychol.* **24**, 319–359. (doi:10.1016/S0065-2601(08)60333-0)
- Kahneman, D. & Miller, D. T. 1986 Norm theory: comparing reality to its alternatives. *Psychol. Rev.* **93**, 136–153. (doi:10.1037/0033-295X.93.2.136)
- Krueger, F., Barbey, A. K. & Grafman, J. In press. The medial prefrontal cortex represents social event knowledge. *Trends Cogn. Sci.*
- Landman, J. 1987 Regret and elation following action and inaction: affective responses to positive versus negative outcomes. *Pers. Soc. Psychol. Bull.* **13**, 524–536. (doi:10.1177/0146167287134009)
- Levy, R. & Goldman-Rakic, P. S. 2000 Segregation of working memory functions within the dorsolateral prefrontal cortex. *Exp. Brain Res.* **133**, 23–32. (doi:10.1007/s002210000397)
- Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F. & Kelley, W. M. 2004 Medial prefrontal activity predicts memory for self. *Cereb. Cortex* **14**, 647–654. (doi:10.1093/cercor/bhh025)
- Mandel, D. R. 2003 Counterfactuals, emotions, and context. *Cogn. Emotion* **17**, 139–159. (doi:10.1080/02699930302275)
- Miller, D. T. & Turnbull, W. 1990 The counterfactual fallacy: confusing what might have been with what ought to have been. *Soc. Justice Res.* **4**, 1–19. (doi:10.1007/BF01048532)
- Milne, E. & Grafman, J. 2001 Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *J. Neurosci.* **21**, RC150.
- Mitchell, J. P., Heatherton, T. F. & Macrae, C. N. 2002 Distinct neural systems subserve person and object knowledge. *Proc. Natl Acad. Sci. USA* **99**, 15 238–15 243. (doi:10.1073/pnas.232395699)
- Moll, J. & de Oliveira-Souza, R. 2007 Moral judgments, emotions, and the utilitarian brain. *Trends Cogn. Sci.* **11**, 319–321. (doi:10.1016/j.tics.2007.06.001)
- Niedenthal, P. M., Tangney, J. P. & Gavanski, I. 1994 "If only I weren't" versus "If only I hadn't": distinguishing shame and guilt in counterfactual thinking. *J. Pers. Soc. Psychol.* **67**, 585–595. (doi:10.1037/0022-3514.67.4.585)
- Ochsner, K. N., Knierim, K., Ludlow, D. H., Hanelin, J., Ramachandran, T., Glover, G. & Mackey, S. C. 2004 Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *J. Cogn. Neurosci.* **16**, 1746–1772. (doi:10.1162/0898929042947829)
- O'Reilly, R. C. & Rudy, J. W. 2000 Computational principles of learning in the neocortex and hippocampus. *Hippocampus* **10**, 389–397. (doi:10.1002/1098-1063(2000)10:4<389::AID-HIPO5>3.0.CO;2-P)
- Palmer, S. E. 1999 *Vision science: photons to phenomenology*. Cambridge, MA: MIT Press.
- Partiot, A., Grafman, J., Sadato, N., Wachs, J. & Hallett, M. 1995 Brain activation during the generation of non-emotional and emotional plans. *Neuroreport* **6**, 1397–1400. (doi:10.1097/00001756-199507100-00009)
- Patterson, R. & Barbey, A. K. In press. Causal simulation theory: an integrative cognitive neuroscience framework for causal reasoning. In *Neural basis of belief systems* (eds J. Grafman & F. Krueger), New York, NY: Psychology Press.
- Rizzolatti, G. 1996 Localization of grasp representations in humans by PET: I. Observation versus execution. *Exp. Brain Res.* **111**, 246–252. (doi:10.1007/BF00227301)
- Roese, N. J. & Olson, J. M. (eds) 1995 *What might have been: the social psychology of counterfactual thinking*, Mahwah, NJ: Erlbaum.
- Roese, N. J., Hur, T. & Pennington, G. L. 1999 Counterfactual thinking and regulatory focus: implications for

- action versus inaction and sufficiency versus necessity. *J. Pers. Soc. Psychol.* **77**, 1109–1120. (doi:10.1037/0022-3514.77.6.1109)
- Schmitz, T. W., Kawahara-Baccus, T. N. & Johnson, S. C. 2004 Metacognitive evaluation, self-relevance, and the right prefrontal cortex. *Neuroimage* **22**, 941–947. (doi:10.1016/j.neuroimage.2004.02.018)
- Sloman, S., Barbey, A. K. & Hotelling, J. In press. A causal model theory of the meaning of cause, enable, and prevent. *Cogn. Sci.*
- Sternberg, R. J. & Gastel, J. 1989 If dancers ate their shoes: inductive reasoning with factual and counterfactual premises. *Mem. Cogn.* **17**, 1–10.
- Thomas, N. J. T. 1999 Are theories of imagery theories of imagination? An active perception approach to conscious mental content. *Cogn. Sci.* **23**, 207–245. (doi:10.1016/S0364-0213(99)00004-X)
- Tulving, E. & Markowitsch, H. J. 1998 Episodic and declarative memory: role of the hippocampus. *Hippocampus* **8**, 198–204. (doi:10.1002/(SICI)1098-1063(1998)8:3<198::AID-HIPO2>3.0.CO;2-G)
- Walton, M. E., Devlin, J. T. & Rushworth, M. F. 2004 Interactions between decision making and performance monitoring within prefrontal cortex. *Nat. Neurosci.* **7**, 1259–1265. (doi:10.1038/nn1339)
- Weingartner, H., Grafman, J., Boutelle, W., Kaye, W. & Martin, P. R. 1983 Forms of memory failure. *Science* **221**, 380–382. (doi:10.1126/science.6867715)
- Williams, J. H. G., Whiten, A., Suddendorf, T. & Perrett, D. I. 2001 Imitation, mirror neurons and autism. *Neurosci. Biobehav. Rev.* **25**, 287–295. (doi:10.1016/S0149-7634(01)00014-8)
- Wood, J. N., Romero, S. G., Makale, M. & Grafman, J. 2003 Category-specific representations of social and nonsocial knowledge in the human prefrontal cortex. *J. Cogn. Neurosci.* **15**, 236–248. (doi:10.1162/089892903321208178)
- Wood, J. N., Tierney, M., Bidwell, L. A. & Grafman, J. 2005a Neural correlates of script event knowledge: a neuropsychological study following prefrontal injury. *Cortex* **6**, 796–804. (doi:10.1016/S0010-9452(08)70298-3)
- Wood, J. N., Knutson, K. M. & Grafman, J. 2005b Psychological structure and neural correlates of event knowledge. *Cereb. Cortex* **15**, 1155–1161. (doi:10.1093/cercor/bhh215)
- Wrosch, C., Bauer, I., Miller, G. E. & Lupien, S. 2007 Regret intensity, diurnal cortisol secretion, and physical health in older individuals: evidence for directional effects and protective factors. *Psychol. Aging* **22**, 319–330. (doi:10.1037/0882-7974.22.2.319)
- Zacks, J. M. & Tversky, B. 2001 Event structure in perception and conception. *Psychol. Bull.* **127**, 3–21. (doi:10.1037/0033-2909.127.1.3)
- Zacks, J. M., Tversky, B. & Iyer, G. 2001 Perceiving, remembering, and communicating structure in events. *J. Exp. Psychol. Gen.* **130**, 29–58. (doi:10.1037/0096-3445.130.1.29)
- Zeki, S. 1993 *A vision of the brain*. Cambridge, MA: Blackwell Scientific Publications, Inc.
- Zeelenberg, M., van der Pligt, J. & Manstead, A. S. R. 1998 Undoing regret on Dutch television: apologizing for interpersonal regrets involving actions or inactions. *Pers. Soc. Psychol. Bull.* **24**, 1113–1119. (doi:10.1177/01461672982410008)