THE ROYAL SOCIETY

*Published online* 26 November 2004

# The emergence of consequential thought: evidence from neuroscience

**Abigail A. Baird*** **and Jonathan A. Fugelsang**

*Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, USA*
(*abigail.baird@dartmouth.edu, jonf@dartmouth.edu*)

The ability to think counterfactually about the consequence of one's actions represents one of the hallmarks of the development of complex reasoning skills. The legal system places a great emphasis on this type of reasoning ability as it directly relates to the degree to which individuals may be judged liable for their actions. In the present paper, we review both behavioural and neuroscientific data exploring the role that counterfactual thinking plays in reasoning about the consequences of one's actions, especially as it pertains to the developing mind of the adolescent. On the basis of the assimilation of both behavioural and neuroscientific data, we propose a brain-based model that provides a theoretical framework for understanding the emergence of counterfactual reasoning ability in the developing mind.

**Keywords:** counterfactual; legal; functional magnetic resonance imaging; neuropsychology; reasoning; brain development

## 1. INTRODUCTION

One's ability to imagine alternative outcomes and understand the consequences of those outcomes is an essential component of human reasoning. Such *counterfactual thinking* typically involves imagining a set of circumstances leading up to an event that may have had a different outcome *if only* a critical preceding event did not take place. For example, consider the case in which an individual runs over a pedestrian while taking an alternative route home to drop off a coworker. Had the coworker not requested a ride home, the driver might not have taken the alternative route, and thus not struck the pedestrian. Given this set of circumstances, an individual can mutate the events preceding the outcome and judge the degree to which certain mutations could change the outcome (e.g. the consequences of their behaviour).

The legal system places great emphasis on this type of reasoning in that it demands that both judges and jurors use counterfactual thinking when determining the degree to which a particular person or event was responsible for a particular outcome (Spellman & Kincannon 2001). This of course raises several critical issues relating to the extent and efficiency with which people are able to reason counterfactually. These issues are relevant not only to the processes by which judges and jurors engage in counterfactual reasoning, but also relate to the extent to which defendants may be seen as liable for their actions.

In the present paper we will focus on this latter issue, with particular emphasis on its relevance to the developing adolescent. Specifically, we delineate: (i) the role that counterfactual thinking plays in reasoning about the consequences of one's actions; (ii) the neural substrates for counterfactual reasoning; and (iii) the limits that brain maturation places on the ability to successfully reason counterfactually by adolescents. Furthermore, we propose a brain-based model that provides a framework for incorporating the cognitive and neural architecture of counterfactual thought, especially as it pertains to the developing mind. Finally, we suggest possible implications of developmental differences in counterfactual thought as they specifically apply to the juvenile justice system.

## 2. PSYCHOLOGICAL ACCOUNTS OF COUNTERFACTUAL REASONING

Given the multifaceted nature of counterfactual reasoning, it is not surprising that researchers from numerous disciplines have used a variety of metrics to measure performance on tasks designed to measure counterfactual reasoning processes. For example, cognitive studies of counterfactual reasoning have typically presented participants with story scenarios that involve a chain of events that may have contributed to a specific outcome and ask participants to either (i) generate counterfactual statements that may follow from a specific outcome (e.g. Guajardo & Turley-Ames 2004), (ii) derive ways in which a specific outcome could be undone by mutating the preceding events (e.g. Spellman & Kincannon 2001), or (iii) make forced choice judgements to inference tasks related to the information contained in the antecedent statements (Thompson & Byrne 2002).

Current psychological work on counterfactual thinking can be traced back to the proposal of the *simulation heuristic* by Kahneman & Tversky (1982). They proposed that people may spontaneously run '*if–then*' simulations when they are presented with information that has a negative outcome or is surprising. Specifically, they found that

* Author for correspondence (abigail.baird@dartmouth.edu).

One contribution of 16 to a Theme Issue 'Law and the brain'.

participants would mentally mutate exceptional antecedent events that preceded a fatal car crash. Recent studies have begun to investigate the extent to which this process of mutation through the generation of *if–then* simulations depends on executive and working memory resources.

This relationship between working memory and complex reasoning has now been firmly established in a number of reasoning domains including analogical reasoning (e.g. Morrison *et al.* 2004), deductive reasoning (e.g. Capon *et al.* 2003) and inductive reasoning (e.g. Reverberi *et al.* 2002). In fact, several prominent theories of reasoning rest on the assumption that both the complexity of the reasoning problems and subsequent demands on working memory relate closely to one's ability to generate alternative representations required for successful reasoning performance. Perhaps the most cited reasoning theory that has recently been applied to counterfactual reasoning (see Byrne 2002; Thompson & Byrne 2002) is the theory of mental models (Johnson-Laird 1983; Johnson-Laird & Byrne 1991). As the name implies, mental models are proposed to be internal mental representations of real or imaginary situations that can be derived from something you directly perceive in the environment. The important thing to consider for the present discourse is that mental models are thought to represent a *possibility* for how different events and relationships between events can be described. According to this theory, deductive reasoning occurs in stages whereby reasoners initially represent and then test tentative conclusions that may follow from the circumstances that they encounter. The final stage of reasoning involves the search and testing of *alternative models* that may be consistent or inconsistent with the initial representations of the scenario. Generating alternative representations, however, requires significant working memory resources and is thus rarely accomplished (e.g. Evans *et al.* 1999). Indeed, recent research has found that there are large individual differences in people's ability to generate alternatives, and subsequently reason effectively (Torrens *et al.* 1999; Newstead *et al.* 2002).

For example, Goldinger *et al.* (2003) examined the degree to which performance on a counterfactual reasoning task was influenced by memory load. Participants were given a series of short stories that contained either a prototypical non-surprising decision (control story), or an unusual surprising decision (counterfactual inducing story). Their task was to take the role of a potential juror and make judgements about compensation for a victim. The critical manipulation was the inclusion of a concurrent memory load task that required participants to hold in memory six bisyllabic words while rendering judgement. This concurrent memory load manipulation was intended to tax working memory and thus leave limited cognitive resources available for the primary counterfactual reasoning task. Participants had also been pre-tested on a simple memory span task in order to control for individual differences in basic working memory capacity.

Results demonstrated that participants were more likely to blame the victim when counterfactual thoughts were easily generated. That is, when participants could easily generate an alternative outcome by mutating the antecedent behaviour of the victim, the victim was judged as more liable for the outcome under question. This was true regardless of baseline memory function. Controlling for individual difference in baseline memory function, however, revealed an interesting result. Specifically, participants with relatively lower memory capacity were increasingly likely to blame the victim as a function of memory load. This occurred presumably because reasoning effectively with counterfactuals requires effortful processing and for individuals with a limited working memory capacity at baseline, such processing capacity is significantly diminished by the addition of a concurrent memory task. These latter findings seem counterintuitive when one considers the prior relationship between ease of alternative generation and victim blame and thus highlight the role of additional components of counterfactual thinking. It is important to note that effective counterfactual reasoning requires not only the generation of appropriate alternative courses of action, but also the ability to sort through and decide which behaviours would result in the appropriate outcome, and which would not. These capacities, as well as their integration, rely heavily on cognitive resources; resources that were not available under the memory load condition for participants with a low capacity at baseline.

## 3. NEURAL SUBSTRATES OF COUNTERFACTUAL REASONING

As described above, the ability to reason counterfactually is highly intertwined with executive function more generally. Because of the paucity of studies directly addressing the brain bases of counterfactual reasoning, it is helpful to interpret the few studies carried out within the context of what is known about the neural substrates of executive function. The term executive function has been used to define complex cognitive processing requiring the flexible coordination of several subprocesses to achieve a particular goal (Funahashi 2001). When these systems break down, behaviour becomes poorly controlled, disjointed and disinhibited (see Elliot (2003) for a review). The structure and function of the frontal lobes are intimately tied to executive processes. Data from both healthy and brain-damaged individuals have provided consistent evidence underscoring the central role that the frontal lobes play in executive function.

Patients with damage to the prefrontal cortex show impaired judgement, organization, planning and decision-making (Stuss & Benson 1984), as well as behavioural disinhibition and impaired intellectual abilities (Luria 2002). Despite the fact that selective aspects of executive function may appear intact in patients with frontal lobe damage, when coordination of a number of functions is required, either in a testing or real life situation, patients with frontal damage are often unable to perform the task (Stuss & Alexander 2000; Elliot 2003). Again, this underscores the significance of the frontal cortex in the generation and coordination of multiple processes that result in appropriate, goal-driven behaviour.

There is a growing body of literature that suggests that one way in which the frontal cortex may enable executive function is by flexibly coordinating with other cortical and subcortical regions. This is not surprising given the fact that the prefrontal cortex has been demonstrated to have reciprocal connections with nearly every part of the brain (see Petrides & Pandya (2002) for a review). In a clever meta-analysis, Duncan & Owen (2000) compared a variety of tasks posited to tap executive function and identified three areas that were reliably active across tasks. The studies examined in the meta-analysis manipulated a particular task demand: response conflict, task novelty, working memory load, memory delay or perceptual difficulty. For each of these demands, five or more studies had assessed the effects of manipulating that demand. Three main clusters were distinguished: dorsal anterior cingulate, a mid-dorsolateral frontal region, and a mid-ventrolateral frontal region. The authors concluded that a common network, involving these three regions, is recruited by diverse cognitive demands. However, they did not rule out the possibility that there may be finer specializations within this network. The investigators suggest that it is possible that the three regions do subserve different functions, but that these functions are sufficiently abstract to be involved in many different complex cognitive tasks.

Similar to the workings of a car, no one would argue that an engine, tyres and steering wheel all serve the same function; however, their most common function is an interdependent and synergistic one, where together they provide transportation. Now, without taking the analogy too far, transportation can be useful to achieve many goals, as can executive function. It is thought that the three regions described above are domain general, but work seamlessly with more domain-specific networks, as is illustrated by tasks requiring semantic memory. Semantic memory is generally considered to be memory for facts, or declarative information. It has been established in the literature that the rich memories that we possess for factual information invoke a distributed network involving both ventral and dorsal streams of processing upon memory retrieval (Thompson-Schill 2003). Here, the degree to which visual, verbal, spatial or sensorimotor networks are recruited during semantic retrieval depends on the degree to which the original encoding episode evoked specialized neural circuitry devoted to domain-specific sensory processing. However, in concert with the recruitment of domain-specific neural circuitry associated with the sensory components of semantic memory, domain general processes in the form of controlled direct memory retrieval act to constrain and deliberate over the outcome of the lower-level sensory processes. This too is the realm of the DLPFC. Succinctly, successful performance on this type of task depends on a minimum of two distinct processes, the ability to generate the information itself and the ability to organize the output in terms of task relevance (i.e. making sure the generated fact fits the imposed category, keeping track of what items have already been said). In summary, the emerging view suggests that multiple brain regions combine with each other in vast numbers of ways, depending on the task requirements and, more generally, on the types of skills that a person, within a specific context, develops (Carpenter *et al.* 2000). Few complex behaviours illustrate executive function better than reasoning.

Consistent with the cognitive demands of everyday reasoning, lesion studies implicate the DLPFC as being essential for everyday reasoning (Shallice & Burgess 1991; Stuss & Alexander 2000). In addition, research by a number of cognitive neuroscientists examining both inductive reasoning (Goel & Dolan 2000; Seger *et al.* 2000) and deductive reasoning (Goel *et al.* 1998; Osherson *et al.* 1998; Parsons & Osherson 2001) using fMRI and positron emission tomography (PET) have pointed to the dominant role of the DLPFC in tasks that demand high-level reasoning. More recently, Goel & Dolan (2004) demonstrated preferential DLPFC activity during an inductive reasoning task (relative to a task that required deductive reasoning). Inductive reasoning is more sensitive to background knowledge rather than logical form, and is essential for knowing which ideas generalize and which do not. The increased activity in DLPFC may thus be because of use of world knowledge in the generation and evaluation of hypotheses (Grafman 2002), which is the basis of inductive reasoning.

Although it is abundantly clear that the prefrontal regions described at length above are involved in many types of reasoning, the additional components required for counterfactual reasoning have remained unclear. In terms of searching for a neural substrate for this type of reasoning, the fronto-striatal system is a likely candidate. There are several convergent lines of evidence that support this notion. In spite of the many functional imaging studies that have examined many types of reasoning, as well as its sub-components, no studies to date have used functional neuroimaging to examine the neural substrates of counterfactual reasoning. Given this, the scarce amount of lesion data becomes increasingly important. Parkinson's disease presents a unique case of a functional lesion. Recently, McNamara *et al.* (2003) have shown reliable deficits in counterfactual reasoning associated with Parkinson's disease.

Parkinson's disease is characterized by rigidity, bradykinesia, gait disorders, and sometimes tremors. The primary pathology involves loss of dopaminergic cells in the substantia nigra and the ventral tegmental area. These two subcortical dopaminergic sites give rise to two projection systems important for motor, affective and cognitive functioning. The nigrostriatal system, primarily implicated in motor functions, originates in the substantia nigra and terminates in the striatum. The meso-limbic–cortical system contributes to cognitive and affective functioning. It originates in the ventral tegmental area and terminates in the ventral striatum, amygdala and frontal lobes, as well as other basal forbrain areas. The degree of nigrostriatal impairment correlates with the degree of motor impairment, while ventral–tegmental–mesocortical impairment correlates positively with the degree of affective and intellectual impairment. The mesocortical dysfunction most probably has a negative impact on prefrontal function.

Previous neuropsychological investigations have consistently reported evidence of prefrontal dysfunction in Parkinson's patients (see McNamara *et al.* (2003) for a review of these findings). In addition to the observed cognitive deficits, the Parkinsonian personality is said to be rigid, stoic and characterized by low novelty seeking. Many patients in the mid to late stages of the disorder are perseverative both on tests of executive function and in their daily tendency to do counterproductive or even dangerous things (for example driving a car, working with power tools, taking on complex construction projects), despite warnings to avoid these dangers. We suggest that impairments in counterfactual reasoning might help explain why these patients fail to learn from past mistakes and thus why they persist in maladaptive or dangerous behaviour.

McNamara *et al.* (2003) asked both Parkinson's patients and age-matched controls to recall an autobiographical event that they perceived as having been negative. Following recall of the event, they were asked if they had any thoughts of how things might have gone differently (e.g. 'what if' or 'only if' type statements). All responses were recorded and the number of distinct counterfactual thoughts were tabulated. The investigators reported that relative to age-matched control subjects Parkinson's patients spontaneously generated fewer counterfactuals than controls, despite showing no differences from controls on a semantic fluency test. Although the authors attribute the reduction in counterfactual generation to frontal dysfunction, it is equally plausible that the deficits arise from the disruption of communication between the basal ganglia and frontal cortices. If the deficits were solely attributable to prefrontal deficits, one might have expected to see more impairment on the semantic fluency task, a task known to rely, in great part, on frontal function. Additional evidence for the role of the basal ganglia in response generation comes from a recent report of deep brain stimulation in Parkinson's patients. Witt *et al.* (2004) surgically stimulated a portion of the basal ganglia (the subthalamic nucleus) in 23 Parkinson's patients, and found that this treatment improved performance on the 'random number generation task', while performance on digit span and verbal fluency tasks were unchanged following treatment. Another group of individuals who show deficits in counterfactual reasoning, but possess a much better prognosis than those with Parkinson's disease, are adolescents.

## 4. ADOLESCENT COGNITION

Adolescence is the period of life between puberty and adulthood. Adolescence begins at the onset of puberty, which technically refers to the time at which an individual is capable of reproduction. Although estimates vary, pubertal onset generally occurs between the ages of 10 and 12 years for girls, and between 13 and 15 years for boys. Once a child is of reproductive age, they have entered adolescence, but are still far from adulthood. Adolescence describes this transitional time, where the individual undergoes major changes in physiological, social, emotional and cognitive functioning that over a period of years enable them to become an adult member of society.

The hallmark of adolescent cognition is the qualitative change that adolescent thinking undergoes. Their thought becomes more abstract, logical and idealistic. Adolescents are more capable of examining their own thoughts, others' thoughts, and what others are thinking about them, and are more likely to interpret and monitor the world around them. What this suggests is that the primary change in adolescent cognition is a dramatic improvement in the ability to think and reason in the abstract. Piaget (1954) believed that adolescents are no longer limited to actual, concrete experiences as anchors for thought. They can create make-believe situations, events that are entirely hypothetical possibilities, or strictly abstract propositions. The primary gain in adolescent cognition is that in addition to being able to generate abstract thought, they are able to reason about the products of their cognition.

There are phases to the emergence of adolescent thought. In the first phase, the increased ability to think hypothetically produces unconstrained thoughts with unlimited possibilities. This early adolescent thought submerges reality (Broughton 1978). During the later phases, the adolescent learns to better regulate their thoughts, measuring the products of their reasoning against experience and imposing monitoring or inhibitory cognitions when appropriate. By late adolescence, many individuals are able to reason in ways that resemble those of adults; however, it is clear that the emergence of this ability depends in great part on experience, and therefore does not appear across all situational domains simultaneously. Simply, adult-like thought is more likely to be used in areas where adolescents have the most experience and knowledge (Carey 1988). During development, adolescents acquire elaborate knowledge through extensive experience and practice in multiple settings (e.g. home, school, sports). The development of expertise in different domains of life bolsters high-level, developmentally mature-looking thought. Experience, and the ability to generalize about it, gives older adolescents two important improvements in reasoning ability. Greater experience, and an improved system for organizing and retrieving the memories of experience, enables the adolescent to recall and apply a greater number of previous experiences to new situations. Additionally, an increased ability to abstract and generalize may allow an adolescent to reason about a situation that they have not directly experienced. Improvements in executive function are largely the result of the synergistic maturation in working memory capacity, selective attention, error detection and inhibition, all of which have been shown to improve with maturational changes in brain structure and function.

## 5. ADOLESCENT BRAIN MATURATION

It is now well established that the overall volume of the human brain does not change dramatically after the age of *ca.* 3 years old (see Thompson *et al.* 2000). Brain maturation, therefore, can be almost entirely attributed to the reorganization and refinement of this relatively fixed space. This process occurs in two distinct ways, myelination and synaptic pruning, each of which directly affects brain functioning. Myelination is the process by which the 'wires' of the brain become insulated. Myelin is a fatty substance that increases the speed with which signals can travel in the brain. It also serves as an index of connectivity within the brain (see Baird *et al.* 2004). Myelinated fibers connect regions of grey matter and enable their communication. Grey matter, or cortex, is where the brain's 'work' is done, it is the light that is lit by the signal delivered via the white matter; it is also where changes in blood oxygenation are measured by fMRI to make inferences about brain function. Synaptic pruning is the process by which the connections within the grey matter are refined, it is believed that the brain follows a strict 'use or lose' policy with regard to grey matter. This process results in a more efficient cortex, and in conjunction with myelination, a more extensively connected cortex.

Perhaps the most consistently reported finding associated with adolescent brain development is the decrease of grey matter and the increase of white matter throughout the cortex, but most significantly within the frontal cortex (see Giedd *et al.* (1999) and Sowell *et al.* (1999) for reviews). The prefrontal cortex is of paramount interest in adolescent development largely because of its well-understood function with regard to cognitive, social and emotional processes in adulthood. The converging evidence of prolonged development and organization of the prefrontal cortex throughout childhood and adolescence (Huttenlocher 1979; Chugani *et al.* 1987; Diamond 1988, 1996) may suggest an important parallel between brain development and cognitive development.

One striking difference regarding the development of the prefrontal cortex relative to other cortical areas is the continuation of synaptic pruning into young adulthood. This decrease in synaptic density during adolescence coincides with the emergence of newly entwined cognitive and emotional phenomena. The secondary process that is likely to be taking place during this time is the fortification of synaptic connections that will remain into adulthood. There has been further speculation that this 'use it or lose it' process may represent the behavioural, and ultimately, the physiological suppressing of immature behaviours that have become obsolete because of the novel demands of adulthood (Casey *et al.* 2000). One can imagine that a response to a particular event in the environment will be potentiated by repeated exposure and subsequent strengthening of the relationship between that event and the generation of the appropriate response. The delayed maturation of this brain region allows the individual to adapt to the particular demands of their unique environment.

Evidence of age-related decreases in grey matter, as a result of synaptic pruning, is borne out in the work of Casey *et al.* (2000) that consistently demonstrated increased volume of cortical activity in younger adolescents who perform less well on tasks of cognitive control and attentional modulation. This pattern of greater brain activity in children relative to adults is suggestive of a gradual decrease in the brain tissue required to perform the task. This observed decrease parallels the observed decrease in frontal grey matter volume, and may result from the elimination of redundant or superfluous synapses (Sowell *et al.* 1999). Many researchers have documented that while there are age-related decreases in grey matter in the prefrontal cortex, the overall cortical volume does not change significantly. Not surprisingly, the cortical volume remains stable because of simultaneous expansion of white matter that may be equally important in terms of functionality (Klingberg *et al.* 1999). The greater volume of frontal white matter observed during adolescence is probably the result of greater axonal myelination. Increases in the speed with which information is processed in the mature brain may reflect this important structural change. Specifically, maturational improvements in frontal connectivity, both within the frontal cortex and with more distal regions (specifically the parietal cortex and basal ganglia, respectively), tracks closely with age-related behavioural improvements in working memory (Olesen *et al.* 2004), as well as response selection and inhibition (Durston *et al.* 2002).

Another consistent set of findings in the neurodevelopmental literature describes changes in basal ganglia structure and connectivity. Sowell *et al.* (1999) reported reductions in the grey matter volume of the basal ganglia, while a number of investigators have reported age-related increases in white matter integrity in the internal capsule, the band of fibres responsible for communication between the thalamus, basal ganglia and cortex (for a review see Schmithorst *et al.* 2002). The structural and functional maturation of the fronto-striatal circuit has been directly related to improvements in performance on the Stroop task (Blumberg *et al.* 2003) and inhibitory control (Durston *et al.* 2002).

In addition to the frontal and basal ganglia, the parietal cortex also undergoes a great deal of developmental change during adolescence. Similar to what has been described above, there have been numerous accounts of parietal grey matter reduction during adolescence (Giedd *et al.* 1999; Sowell *et al.* 1999; Thompson *et al.* 2000). These changes in the parietal cortex track closely with similar reductions observed in the frontal cortex (Sowell *et al.* 2003). Additionally, the white matter tracks connecting frontal and parietal cortices become increasingly embellished during adolescence as evidenced by both magnetic resonance (MR) studies (Olesen *et al.* 2004) and electroencephalogram (EEG) studies (Thatcher 1994). Finally, the maturation of fronto-parietal connectivity has been found to be highly correlated with improvements in working memory (Olesen *et al.* 2004). In addition to working memory, the observed refinements in the parietal cortex are also likely to contribute to improvements in mental imagery as well as representations of one's body in space. It is conceivable that one role that the mature parietal cortex plays is allowing an individual to 'try on' a particular experience (i.e. envision themselves engaging in an imagined activity). In the case of counterfactual reasoning, it is plausible that the 'what ifs' generated by the basal ganglia are sent to the prefrontal cortex for logical approval, and following initial approval are forwarded to the parietal cortex to 'see what it might look like', to check the feasibility of the behaviour for the individual. For example, this has become a rather common practice for individuals using mental imagery to improve their golf swing, a method that has recently been shown to produce activity in both prefrontal and parietal cortex (Ross *et al.* 2003). A hypothesized graphical depiction of the brain-based networks subserving counterfactual thought, their interrelations and developmental trajectory is depicted in figure 1.

## 6. IMPLICATIONS FOR LEGAL THOUGHT

What does the development of counterfactual reasoning mean for the justice system? One direct implication of this model is that young adolescents may lack the neural hardware to generate behavioural alternatives in situations demanding a response. For example, adolescents are more likely than most adults to engage in risk-taking behaviour. While there are a myriad of theories about why this is the case (see Spear (2002) for an extensive review), one reason for increased risk taking in adolescents might be their inability to generate alternatives and potential outcomes prior to the initiation of behaviour. More specifically, a great number of adults think about driving their cars at excessive speeds, and while some adults do engage in this behaviour, adults are more likely to also envision a number of counterfactual scenarios that vary in their desirability. This is an important component of appreciating potential consequences of actions.

As stated in § 1, the legal system places great emphasis on the ability to appreciate the consequences of one's actions. Additionally, jurors are more likely to assign greater levels of culpability if they believe that a specific outcome under

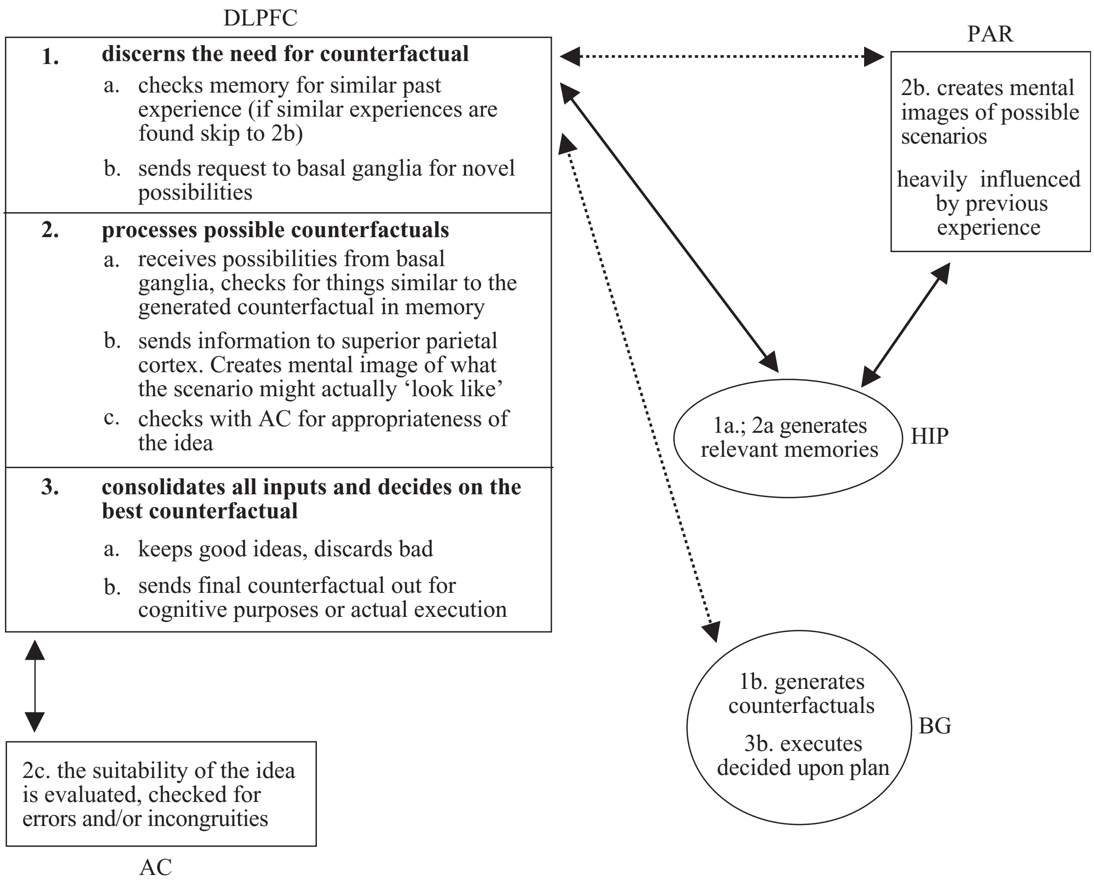


Figure 1. A proposed brain-based model for the generation of counterfactual reasoning. AC, anterior cingulate; BG, basal ganglia; HIPP, hippocampus; PAR, parietal cortex. The dashed lines indicate projections that mature during adolescence.

deliberation would not have occurred 'if only' the action carried out by an individual or organization ensued an alternative action preceding the outcome. As previously stated, adolescents are much more likely to reason abstractly about situations where they have had some previous experience; however, it is the interaction of continued experience and refinements in the adolescent brain that enable the emergence of counterfactual reasoning, as well as the appreciation of consequences, in the absence of actual experience. What the evidence presented in this review suggests is that it may be physically impossible for adolescents to engage in counterfactual reasoning, and as a result of this are often unable to effectively foresee the possible consequences of their actions.

On 19 July 2004, a large number of organizations including the leading American medical, religious and legal institutions, child and victim advocate groups and representatives from nearly 50 countries, along with prominent individuals, including Nobel Peace Prize Laureates and former US Diplomats, filed amicus curiae briefs calling for an end to the juvenile death penalty. 'In their friend-of-the-court briefs, the groups state that the juvenile death penalty violates evolving standards of decency, that it serves no legitimate purpose and is excessive in light of emerging evidence showing the limited capabilities of juveniles, and that the practice is almost universally rejected by the international community' (see http://www.deathpenaltyinfo.org).

Both the American Medical Association and the American Psychological Association have submitted briefs to the United States Supreme Court, reviewing deficiencies in brain structure, function and concomitant behaviour in adolescents relative to adults. Accordingly, these deficiencies warrant exclusion from the death penalty and violate the Eighth Amendment of the United States Constitution, which states that 'excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted'. The spirit of this logic is captured in the case of Thompson versus Oklahoma, where the Supreme Court prohibited the execution of juveniles whose crimes were committed prior to their sixteenth birthday. Justice Stevens emphasized the relative immaturity of adolescent cognition:

> Less culpability should attach to a crime committed by a juvenile than to a comparable crime committed by an adult. The basis of this conclusion is too obvious to require extensive explanation. Inexperience, less intelligence and less education make a teenager less able to evaluate the consequences of his or her conduct while at the same time he or she is more apt to be motivated by mere emotion or peer pressure than is an adult. The reasons that juveniles are not trusted with the privileges and responsibilities of an adult also explain why their irresponsible conduct is not as morally reprehensible as that of an adult.
> (Thompson versus Oklahoma 1998, 487, U.S. 815, p. 835).

The final sentence of Justice Steven's remarks provides the most irony with regard to juvenile justice. In the state of Texas, where more than half of all executions for crimes committed as juveniles have occurred, one cannot attend an R-rated movie unless they are 17 years old. Additionally, an individual cannot purchase a lottery ticket, and perhaps most importantly, serve on a jury, until the age of 18. For juvenile defendants, a jury made up of 15, 16 and 17 year olds is clearly called for by the sixth amendment of the United States Constitution, which guarantees a defendant the right to a trial by a jury of their peers. The idea of a teenage jury violates not only common sense, but also the laws that mandate individuals be a minimum of 18 years old to serve on a jury. Therein lies a bewildering conflict.

While an infant is learning to walk, there may be days when they pull themselves up on the furniture or their parents, they may also balance on their feet without holding onto anything. As anyone who has witnessed an infant learning to walk can report, the first few steps are usually followed by a tumble and sporadic reattempts. They may take their first steps one day, and then not walk again for several days following this initial foray. We acknowledge that this process is nonlinear and a result of the interaction of both an immature brain and lack of experience. Within the developing adolescent, the emergence of adult levels of reasoning is no different. An adolescent may demonstrate an adult-like ability to reason abstractly, and act in accordance with this advanced cognition on Monday, but behave impulsively and irrationally on Thursday. What these two examples have in common is the idea that the appearance of a behaviour does not indicate its permanence. Adolescence is an awkward time, both in terms of movement and thinking, during which the individual becomes increasingly coordinated.

Developmental neuroscience has yet to conclusively offer a chronological age at which we can be absolutely certain that an individual's brain structure and function have become fully mature. A great deal of work remains to be carried out to answer this question convincingly. Until such time, it is of paramount importance that researchers, practitioners and members of the legal community continue to work together to ensure the well being of both our society and its children.

## REFERENCES


Baird, A. A., Colvin, M. K., VanHorn, J., Inati, S. & Gazzaniga, M. S. 2004 Functional connectivity: integrating behavioral, DTI and fMRI datasets. *J. Cogn. Neurosci.* (In the press.)

Blumberg, H. P., Kaufman, J., Martin, A., Whiteman, R., Zhang, J. H., Gore, J. C., Charney, D. S., Krystal, J. H. & Peterson, B. S. 2003 Amygdala and hippocampal volumes in adolescents and adults with bipolar disorder. *Arch. Gen. Psychiatry* **60**, 1201–1208.

Broughton, J. M. 1978 Development of concepts of self, mind, reality, and knowledge. In *Social cognition* (ed. W. Damon), pp. 75–100. San Francisco, CA: Josey-Bass.

Byrne, R. M. J. 2002 Mental models and counterfactual thinking. *Trends Cogn. Sci.* **6**, 405–445.

Capon, A., Handley, S. & Dennis, I. 2003 Working memory and reasoning: an individual differences perspective. *Thinking Reasoning* **9**, 203–244.

Carey, S. 1988 Are children fundamentally different kinds of thinkers and learners than adults? In *Cognitive development to adolescence* (ed. K. Richardson & S. Sheldon), pp. 105–138. Hillsdale, NJ: Earlbaum.

Carpenter, P. A., Just, M. A. & Reichle, E. D. 2000 Memory and executive function: evidence from neuroimaging. *Curr. Opin. Neurobiol.* **10**, 195–199.

Casey, B. J., Giedd, J. N. & Thomas, K. M. 2000 Structural and functional brain development and its relation to cognitive development. *Biol. Psychol.* **54**, 241–257.

Chugani, H. T., Phelps, M. E. & Mazziotta, J. C. 1987 Positron emission tomography study of human brain functional development. *Ann. Neurol.* **22**, 487–497.

Diamond, A. 1988 Abilities and neural mechanisms underlying AB performance. *Child Dev.* **59**, 523–527.

Diamond, A. 1996 Evidence for the importance of dopamine for prefrontal cortex functions early in life. *Phil. Trans. R. Soc. Lond.* B **351**, 1483–1493.

Duncan, J. & Owen, A. M. 2000 Common regions of the human frontal lobes recruited by diverse cognitive demands. *Trends Neurosci.* **23**, 475–483.

Durston, S., Thomas, K. M., Worden, M. S., Yang, Y. & Casey, B. J. 2002 The effect of preceding context on inhibition: an event-related fMRI study. *Neuroimage* **16**, 449–453.

Elliot, R. 2003 Executive functions and their disorders. *Br. Med. Bull.* **65**, 49–55.

Evans, J. S. B. T., Handley, S. J., Harper, C. N. J. & Johnson-Laird, P. N. 1999 Reasoning about necessity and possibility: a test of the mental model theory of deduction. *J. Exp. Psychol. Learn.* **25**, 1495–1513.

Funahashi, S. 2001 Neuronal mechanisms of executive control by the prefrontal cortex. *Neurosci. Res.* **39**, 147–165.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C. & Rapoport, J. L. 1999 Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neurosci.* **10**, 861–863.

Goel, V. & Dolan, R. J. 2000 Anatomical segregation of component processes in an inductive inference task. *J. Cogn. Neurosci.* **12**, 110–119.

Goel, V. & Dolan, R. J. 2004 Differential involvement of left prefrontal cortex in inductive and deductive reasoning. *Cognition* **93**, B109–B121.

Goel, V., Gold, B., Kapur, S. & Houle, S. 1998 Neuroanatomical correlates of human reasoning. *J. Cogn. Neurosci.* **10**, 293–302.

Goldinger, S. D., Kleider, H. M., Azuma, T. & Beike, D. R. 2003 Blaming the victim under memory load. *Psychol. Sci.* **14**, 81–85.

Grafman, J. 2002 The structured event complex and the human prefrontal cortex. In *The frontal lobes* (ed. D. T. Stuss & R. T. Knight), pp. 292–310. Oxford University Press.

Guajardo, N. R. & Turley-Ames, K. J. 2004 Preschoolers' generation of different types of counterfactual statements and theory of mind understanding. *Cogn. Dev.* **19**, 53–80.

Huttenlocher, P. R. 1979 Synaptic density in human frontal cortex: developmental changes and effects of aging. *Brain Res.* **163**, 195–205.

Johnson-Laird, P. N. 1983 *Mental models*. Cambridge University Press.

Johnson-Laird, P. N. & Byrne, R. M. J. 1991 *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kahneman, D. & Tversky, A. 1982 The simulation heuristic. In *Judgement under uncertainty: heuristics and biases* (ed. D. Kahneman, P. Slovic & A. Tversky), pp. 201–210. Cambridge University Press.

Klingberg, T., Vaidya, C. J., Gabrieli, J. D., Moseley, M. E. & Hedehus, M. 1999 Myelination and organization of the frontal white matter in children: a diffusion tensor MRI study. *Neuroreport* **10**, 2817–2821.

Luria, A. R. 2002 Frontal lobe syndromes. In *Handbook of clinical neurology,* vol. 2 (ed. P. J. Pinken & G. W. Bruyn), pp. 725–757. Amsterdam: North Holland.

McNamara, P., Durso, R., Brown, A. & Lynch, A. 2003 Counterfactual cognitive deficit in persons with Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **74**, 1065–1070.

Morrison, R. G., Krawczyk, D. C., Holyoak, K. J., Hummel, J. E., Chow, T. W., Miller, B. L. & Knowlton, B. J. 2004 A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. *J. Cogn. Neurosci.* **16**, 260–271.

Newstead, S. E., Thompson, V. A. & Handley, S. J. 2002 Generating alternatives: a key component in human reasoning? *Mem. Cogn.* **30**, 129–137.

Olesen, P. J., Westerberg, H. & Klingberg, T. 2004 Increased prefrontal and parietal activity after training of working memory. *Nature Neurosci.* **7**, 75–79.

Osherson, D., Perani, D., Cappa, S., Schnur, T., Grassi, F. & Fazio, F. 1998 Distinct brain loci in deductive versus probabilistic reasoning. *Neuropsychologia* **36**, 369–376.

Parsons, L. M. & Osherson, D. 2001 New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. *Cereb. Cortex* **11**, 954–965.

Petrides, M. & Pandya, D. N. 2002 Association pathways of the prefrontal cortex and functional observations. In *Principles of frontal lobe function* (ed. D. T. Stuss & R. T. Knight), pp. 31–51. Oxford University Press.

Piaget, J. 1954 The construction of reality in the child. (Translated by M. Cook.) New York: Basic Books.

Reverberi, C., Lavaroni, A., Gigli, G. L., Skrap, M. & Shallice, T. 2002 Inductive inferences and the frontal lobes. *Cortex* **38**, 899–902.

Ross, J. S., Tkach, J., Ruggieri, P. M., Lieber, M. & Lapresto, E. 2003 The mind's eye: functional MR imaging evaluation of golf motor imagery. *Am. J. Neuroradiol.* **24**, 1036–1044.

Schmithorst, V. J., Wilke, M., Dardzinski, B. J. & Holland, S. K. 2002 Correlation of white matter diffusivity and anisotropy with age during childhood and adolescence: a cross-sectional diffusion-tensor MR imaging study. *Radiology* **222**, 212–218.

Seger, C., Poldrack, R., Prabhakaran, V., Zhao, M., Glover, G. & Gabrieli, J. 2000 Hemispheric asymmetries and individual differences in visual concept learning as measured by functional MRI. *Neuropsychologia* **38**, 1316–1324.

Shallice, T. & Burgess, P. W. 1991 Deficits in strategy application following frontal lobe damage in man. *Brain* **114**, 727–741.

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L. & Toga, A. W. 1999 *In vivo* evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neurosci.* **2**, 859–860.

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L. & Toga, A. W. 2003 Mapping cortical change across the human life span. *Nature Neurosci.* **6**, 309–315.

Spear, L. P. 2002 The adolescent brain and age-related behavioral manifestations. *Neurosci. Behav. Rev.* **24**, 417–463.

Spellman, B. A. & Kincannon, A. 2001 The relation between counterfactual ('but for') and causal reasoning: experimental findings and implications for jurors' decisions. *Law Contemp. Probl.* **64**, 241–264.

Stuss, D. T. & Alexander, M. P. 2000 Executive functions and the frontal lobes: a conceptual view. *Psychol. Res.* **63**, 289–298.

Stuss, D. T. & Benson, D. F. 1984 Neuropsychological studies of the frontal lobes. *Psychol. Bull.* **95**, 3–28.

Thatcher, R. W. 1994 Cyclic cortical reorganization: origins of human cortical development. In *Human behavior and the developing brain* (ed. G. Dawson & K. Fischer), pp. 232–266. New York: Guilford Press.

Thompson, P. M., Giedd, J. N., Woods, R. P., MacDonald, D., Evans, A. C. & Toga, A. W. 2000 Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature* **404**, 190–193.

Thompson, V. A. & Byrne, R. M. J. 2002 Reasoning counterfactually: making inferences about things that didn't happen. *J. Exp. Psychol. Learn.* **28**, 1154–1170.

Thompson-Schill, S. L. 2003 Neuroimaging studies of semantic memory: inferring 'how' from 'where'. *Neuropsychologia* **41**, 280–292.

Torrens, D., Thompson, V. A. & Cramer, K. M. 1999 Individual differences and the belief bias effect: mental models, logical necessity, and abstract reasoning. *Thinking Reasoning* **5**, 1–28.

Witt, K., Pulkowski, U., Herzog, J., Lorenz, D., Hamel, W., Deuschl, G. & Krack, P. 2004 Deep brain stimulation of the subthalamic nucleus improves cognitive flexibility but impairs response inhibition in Parkinson disease. *Arch. Neurol.* **61**, 697–700.

## GLOSSARY

DLPFC: dorsolateral prefrontal cortex
fMRI: functional magnetic resonance imaging