# Medial prefrontal and anterior cingulate cortical thickness predicts shared individual differences in self-generated thought and temporal discounting

Boris C. Bernhardt [a,*,1], Jonathan Smallwood [a,b,1], Anita Tusche [a], Florence J.M. Ruby [a,b], Haakon G. Engen [a], Nikolaus Steinbeis [a], Tania Singer [a]

[a] Department of Social Neuroscience, Max Planck Institute for Human Cognitive Brain Sciences, Leipzig, Germany
[b] Department of Psychology, University of York, York, UK

## ARTICLE INFO

## ABSTRACT

When deprived of compelling perceptual input, the mind is often occupied with thoughts unrelated to the immediate environment. Previous behavioral research has shown that this self-generated task-unrelated thought (TUT), especially under non-demanding conditions, relates to cognitive capacities such as creativity, planning, and reduced temporal discounting. Despite the frequency and importance of this type of cognition, little is known about its structural brain basis. Using MRI-based cortical thickness measures in 37 participants, we were able to show that individuals with a higher tendency to engage in TUT under low-demanding conditions (but not under high-demanding conditions) show an increased thickness of medial prefrontal cortex (mPFC) and anterior/midcingulate cortex. Thickness of these regions also related to less temporal discounting (TD) of monetary rewards in an economic task, indicative of more patient decision-making. The findings of a shared structural substrate in mPFC and anterior/midcingulate cortex underlying both TUT and TD suggest an important role of these brain regions in supporting the self-generation of information that is unrelated to the immediate environment and which may be adaptive in nature.

© 2013 Elsevier Inc. All rights reserved.

## Introduction

Despite continually receiving perceptual input, the mind spends significant amounts of time generating cognition without a basis in immediate sensory information. Such self-generated thought can derail performance in high-demanding tasks (McVay and Kane, 2011) and is a correlate of unhappiness (Killingsworth and Gilbert, 2010; Smallwood et al., 2009a). The capacity to self-generate thought without recourse to immediate perceptual input can, however, be adaptive. By allowing the individual to focus on information represented in memory, the capacity to self-generate thought can help people make progress on personally relevant goals in a patient and creative manner (Schacter et al., 2012; Smallwood and Andrews-Hanna, 2013).

One method of investigating self-generated thought is to examine individuals' tendencies to engage in thoughts irrelevant to a task being performed, a phenomenon known as task-unrelated thought (TUT). Importantly, the role of TUT can be readily studied in relationship to the difficulty of the ongoing task. As it is hypothesized to depend on the production of mental content unrelated to perception and maintenance of this information in attention, the occurrence of TUT is generally reduced during tasks whose performance depends heavily on continuous attention to external input [for a review, see (Smallwood, 2013)].

Although there are well documented negative consequence of TUT during perceptually challenging tasks, its occurrence under non-demanding conditions is often linked to more beneficial psychological outcomes [for a review, see (Smallwood and Andrews-Hanna, 2013)]. It has been suggested that TUT, especially under low-demanding situations, can help an individual to devote cognitive resources away from present constraints and to consider temporally distant goals (Schooler et al., 2011). Consistent with this hypothesis, it has been shown that individuals who generate TUT under low-demanding circumstances may consolidate self-relevant information more effectively (Smallwood et al., 2011b) and may generate more creative solutions to problems (Baird et al., 2012). Recently, we found that individuals who display more TUT during low demanding situations may engage in less temporal discounting (TD) during economic decision-making tasks (Smallwood et al., 2012b), an indicator of more patient economic decision making (Kable and Glimcher, 2007; Meier and Sprenger, 2012). It is possible that this link between TUT and TD relates to our general ability to decouple from the present input and to engage in self-generated thought (Frith and Frith, 1999, 2003; Mitchell, 2009; Schacter et al., 2008; Smallwood et al., 2012b).

* Corresponding author at: Department of Social Neuroscience, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04301 Leipzig, Germany. Fax: +49 341 9940 2356.
  E-mail address: bernhardt@cbs.mpg.de (B.C. Bernhardt).
[1] Both authors contributed equally to this work.

Previous functional MRI studies have identified that mental processes that require a form of decoupling from perceptual input, such as TUT (Christoff et al., 2009; Mason et al., 2007), but also mentalizing (Amodio and Frith, 2006; Frith and Frith, 1999, 2003; Mitchell, 2009), future thinking (Addis et al., 2007; Buckner and Carroll, 2007; Schacter et al., 2012), meta-cognition (Baird et al., 2013), or autobiographical recall (Hassabis and Maguire, 2007; Huijbers et al., 2011) relate to processes occurring in medial prefrontal cortex, mPFC. Several previous functional studies on TD have also shown that more patient decision-making relates to activity in mPFC (Kable and Glimcher, 2007). Yet, while these functional imaging studies have provided insights into the functional neural substrates involved in self-generated thought, very little is known about the structural basis of inter-individual differences in these capacities.

Recent advances in MRI-based post-processing techniques have allowed the study of structural substrates underlying individual differences in cognitive abilities. In particular, MRI-based cortical thickness measurements have been established as a direct and biologically meaningful marker to quantify continuous structural variations with respect to the anatomy of the folded cortical surface (Dale et al., 1999; Kim et al., 2005; MacDonald et al., 2000; Thompson et al., 2004). Advanced image processing methods enable automatic and reliable measurement of cortical thickness by calculating the distance between the gray matter and white matter surfaces across the entire cortical mantle. While the exact functional substrates of cortical thickness variations remain to be investigated, previous studies have demonstrated links with general intelligence (Colom et al., 2013; Karama et al., 2011), behavioral and trait markers of impulsivity (Schilling et al., 2013; Steinbeis et al., 2012) and strategic behavior (Steinbeis et al., 2012). Building on this work, the current work assessed structural markers of two types of decoupled cognitive processes (TUT and reduced delay discounting) to provide insight into the biological basis of this common, and important, element of human cognition.

We applied MRI-based cortical thickness measures to investigate structural brain substrates that underlie individual differences in two different elements of self-generated thought. We were interested in the study of correlates of TUT, particularly under low demand conditions, and TD as a measure of considering future rewards in the presence of immediate economic incentives. In 37 healthy participants, we examined TUT during a high-demanding working memory task ($TUT_{WM}$) and during a low-demanding choice reaction time task ($TUT_{CRT}$). Based on prior research (Smallwood et al., 2009b, 2011a, 2012b), we expected higher $TUT_{CRT}$ than $TUT_{WM}$. Moreover, given functional MRI findings suggesting a domain-general role of mPFC in self-generated thought processes decoupled from perception, such as TUT and TD (Amodio and Frith, 2006; Christoff et al., 2009; Frith and Frith, 1999; Huijbers et al., 2011; Mason et al., 2007; Mitchell, 2009; Schacter et al., 2008), we expected to find shared structural substrates of $TUT_{CRT}$ and TD primarily in these areas.

## Material and methods

### Participants

We studied 37 healthy volunteers (18 females, aged between 21 and 38 years, mean ± SD age = 27.0 ± 3.8 years, 36 right-handed, one ambidextrous; with 10–13 years of high-school education completed). None of the participants participated in our previous behavioral study (Smallwood et al., 2012b). Participants were German native speakers, free of psychiatric or neurological history, and had normal or corrected-to-normal vision. Written and informed consent was obtained from all participants, and the local ethics committee approved our study.

### Behavioral session

The behavioral session lasted approximately 2 h. Tasks were counterbalanced across subjects. Participants were allowed in taking short breaks between the blocks if desired. One block included two tasks in which we measured TUT; another block consisted of temporal discounting (TD) measurements. A number of other tasks were measured during the third block and will be described elsewhere. Participants gave written and informed consent before the beginning of the experiment and were remunerated at least 16 € for their participation (8 € per hour of participation plus an additional reward depending on their performance during the TD task). E-prime 2 was used for stimulus presentation (Psychology Software Tools Inc., Sharpsburg, PA).

### Task-unrelated thought (TUT)

To measure TUT, participants performed two tasks (a choice reaction time task, CRT; and a 1-back working-memory task, WM). Each task lasted approximately 7 min and participants could take a short break between them if desired. Both CRT and WM task were counterbalanced across participants. During the CRT (low-demand condition), participants observed a sequence of black digits on a computer screen while waiting for a target (a colored digit) to appear, at which point they had to indicate the parity of this target (odd or even) with a button press. In the WM (high-demand condition), participants were exposed to the same sequence of black digits, and were intermittently probed with a colored question mark ('?'). When the question mark was presented, participants had to indicate the parity of the previous digit. Because the occurrence of the colored question mark is randomly determined, this task requires participants to encode and retain in memory the parity of each non-colored number. This task, therefore, requires continuous attention to external information.

In both tasks, black digits were presented for 1000 ms and colored stimuli were presented for 2000 ms against white background. Stimuli were separated by a fixation cross of random duration (2200, 2800, 3200 or 4400 ms). Targets (or question marks) and non-targets were presented with a ratio of approximately 1/6.

Participants were told that they would be asked to report the contents of their thoughts during the experiment. During both the CRT and WM, TUT was recorded using thought probe sampling (mean ± SE number of probes: CRT = 7.10 ± 0.2; WM = 7.07 ± 0.2). Intermittently throughout the tasks, participants were asked to report whether their thoughts were related or unrelated to the task using a Likert scale (1 to 9), with a greater score indicating more off-task focus. For an example of this method, see our previously published behavioral analysis (Smallwood et al., 2012b).

Although it is possible that experience sampling as performed in the current work may lead to higher reports of off-task thinking, studies using retrospective measures recorded at the end of the session confirmed that greater off-task thought occurs in the CRT than in the WM task (Baird et al., 2012).

### Temporal discounting (TD)

As in our previous behavioral experiment (Smallwood et al., 2012b), participants performed a TD task. The TD task lasted approximately 10 min and trial presentation was subject-paced. On each trial, participants had to make a choice between two options: a smaller immediate reward and a larger later reward. Immediate and later options were presented left or right from a central fixation cross, in a counterbalanced order to avoid repetitive behavior.

The immediate reward was fixed at 10 €. The amount of the later reward and the delay at which it could be received varied across trials. Seven different delay periods (ranging from 1 to 180 days) and seven amounts of money (ranging from 12 to 50 €) were used, leading to 49 different trial types. Following the paradigm of Kable and Glimcher (2007), each trial type was presented four times, yielding 196 trials in total. Each trial was separated by a fixation cross (1000 ms) and trial order was randomized. In this paradigm, TD is operationalized as a reduced preference for future rewards, especially at longer delay periods.

In addition to receiving 16 € for their overall participation, subjects received an additional remuneration depending on the choices made

during the TD task. This manipulation was applied so that decisions made during the task were not only hypothetical but also could lead to a real gain (Kable and Glimcher, 2007). Prior to starting the task, participants were informed that one trial from the TD task would be randomly selected and that they would receive 20% of the amount that they had chosen on that trial, at the indicated time point. In reality, all participants received the additional remuneration at the end of the experiment, regardless of the specified delay, to ease the payment process. During debriefing, no participant indicated that they expected this was the case; moreover, subject-scheduling order was uncorrelated with the overall discounting rate (r = −0.06, p > 0.6).

*Structural MRI*

*MRI acquisition*

MRI data were acquired on a 3T Siemens Verio scanner (Siemens Medical Systems, Erlangen), approximately two months prior to the behavioral experiment. Structural images were acquired using a MPRAGE T1-weighted sequence (TR = 2300 ms; TE = 2.98 ms; TI = 900; flip angle = 9°; 176 sagittal slices; matrix size = 240 × 256; FOV = 240 × 256 mm; slice thickness = 1 mm; ipat = 2), yielding a final voxel size of $1.0 \times 1.0 \times 1.0$ mm.

*Cortical thickness measurements*

FreeSurfer (5.1.0; http://surfer.nmr.mgh.harvard.edu) was used to generate models of the cortical surface and to model cortical thickness from the T1-weighted images. Previous work has validated FreeSurfer by comparing it with histological analysis (Rosas et al., 2002) and manual measurements (Kuperberg et al., 2003). The processing steps have been described in detail elsewhere (Dale et al., 1999; Fischl et al., 1999; Han et al., 2006). In brief, MRI data first underwent a series of preprocessing steps that involved intensity normalization, removal of non-brain tissue, tissue classification, and surface extraction. Following surface extraction, sulcal and gyral features across individual subjects were aligned by morphing each subject's brain to an average spherical representation, fsaverage that allows for accurate matching of cortical thickness measurement locations among participants, while minimizing metric distortion. The entire cortex in each subject was visually inspected and segmentation inaccuracies were manually corrected. Cortical thickness was calculated as the closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the tessellated surface. For whole-brain analysis, thickness data were smoothed on the tessellated surfaces using a 20 mm full-with-at-half-maximum Gaussian kernel prior to statistical analysis. Selecting a surface-based kernel reduces measurement noise but preserves the capacity for anatomical localization, as it respects cortical topological features (Lerch and Evans, 2005).

*Statistical analysis*

*Behavioral analysis*

Task-unrelated thought (TUT) was defined as the average number of probes that got a score of four or more on the thought probes (Smallwood et al., 2012b). We compared TUT during the CRT condition ($TUT_{CRT}$) to TUT during the WM condition ($TUT_{WM}$) using paired t-tests.

The participants' preference for future rewards during the TD task was defined as the likelihood of choosing the future reward. Data were analyzed using Analysis of Variance (ANOVA), with 7 levels each reflecting different delay periods. To also assess the relationship between TD and TUT, participants' reports of $TUT_{CRT}$ and $TUT_{WM}$ were both included as continuous between-participant covariates.

*Structural MRI analysis*

Cortical thickness data were analyzed using the SurfStat toolbox for Matlab [http://www.math.mcgill.ca/keith/surfstat, (Worsley et al., 2009)]. We used linear regression models to assess effects of $TUT_{CRT}$,

$TUT_{WM}$, and TD on cortical thickness at each vertex. As in previous work (Bernhardt et al., 2010, in press-a,b; Steinbeis et al., 2012), findings from our surface-based analysis were controlled for multiple comparison using random field theory for non-isotropic images (Worsley et al., 1999). This controlled the chance of reporting a family-wise error (FWE) to p < 0.05.

## Results

*Behavioral findings*

a) TUT. In line with previous findings (Baird et al., 2011; Smallwood et al., 2012b), participants reported higher frequency of TUT in the thought probes for the CRT than for the WM task (paired t-test, t = 3.45, p < 0.001, Fig. 1A). The findings are indicative of a lower focus on the task in hand in the CRT (mean ± SE = 0.68 ± 0.04; range = 0–1) than in the WM task (mean ± SE = 0.50 ± 0.05; range = 0–1).

b) TUT and TD. ANOVAs of TD data with $TUT_{CRT}$ and $TUT_{WM}$ as continuous between-participant covariates yielded an effect of delay on preference (F = 16.88, p < 0.001, Greenhouse–Geisser corrected, see Fig. 1B), indicating a decline in preference for a future reward with increasing delay. In addition, a $TUT_{CRT}$ x delay interaction (F = 3.73, p < 0.05, Greenhouse–Geisser corrected) indicated that increasing reports of $TUT_{CRT}$ were associated with differences in the likelihood of choosing the distant option at the distant time frames. To further understand the relation between $TUT_{CRT}$ and TD, we applied principal components analysis (PCA) to the TD data, and extracted two orthogonal components, which together explained 90% of variance (Fig. 1C). While the score of the first component remained relatively constant over different delays (suggesting that it may closely relate to the overall mean choice), scores of the second component captured the main decline effect with longer periods. Likewise, the second, but not the first component, negatively correlated with $TUT_{CRT}$ (r = −0.38, p < 0.05, Figs. 1D and E). These findings indicate that a participants' likelihood for the self-generation of TUT under non-demanding conditions was associated with a less steep decline in economic rewards at longer intervals. This relation held true when controlling for additive effects of $TUT_{WM}$ in the same statistical model (r = −0.47, p < 0.005). Importantly, no component correlated with $TUT_{WM}$ (all r < 0.1, p > 0.5). These data replicate previous findings showing that subjects with higher reports of $TUT_{CRT}$ have a smaller decline in preference for distant future rewards than subjects with lower reports of $TUT_{CRT}$ (Smallwood et al., 2012b).

To further understand the specific effects, we conducted a linear regression between $TUT_{CRT}$ and the likelihood of selecting a distant reward, controlling for effects of $TUT_{WM}$. This yielded a model that was a significant [F = 3.76, p < 0.005], indicating that $TUT_{CRT}$ is associated with choosing the distant option.

Following previous work (Kable and Glimcher, 2007; Smallwood et al., 2012b), the current study calculated the k-parameter that describes the hyperbolic decline of the preference for the future reward for each individual. For a given individual, we calculated for each combination of a delay D (ranging from 1 to 180 days, in 7 levels) and future reward magnitude M (ranging from 12 to 50 Euros, in 7 levels) the subjective value S, which was defined as the overall likelihood of selecting the future reward over the smaller but immediate 10 Euro reward across the four identical trials in a given D–M combination. Next, we calculated k using the formula $S = 1 / (1 + kD)$. Fitting of k was then carried out across all S associated to each of 7 delays. In our sample, however, we did not find a noteworthy relationship between the k parameter of discounting as an alternative measure of exponential decay and $TUT_{CRT}$ (r < 0.15, p > 0.4). This may have been because k in our sample had a more skewed and spiky distribution (skew = 2.71;
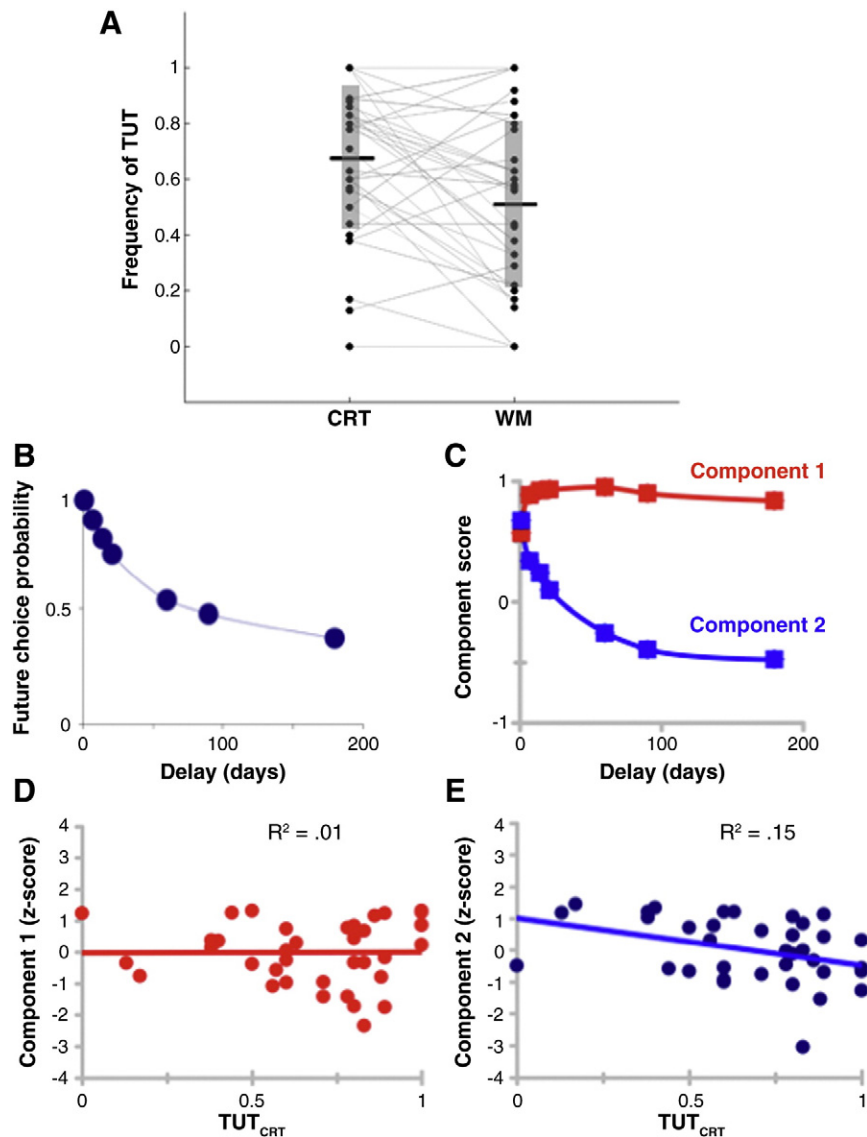
**Fig. 1.** Task-unrelated thought during low-demanding choice reaction time task (TUT$_{CRT}$) and high-demanding working memory (TUT$_{WM}$) task and relationship to temporal discounting (TD). A) Proportions of TUT$_{CRT}$ and TUT$_{WM}$ reported in the current study. Gray lines interconnect measurements from individual subjects. Mean and SD of the proportions of TUT are also indicated. B) Raw TD data indicating the decay of preference for a future reward across increasing delays (from 1 to 180 days) to obtain the reward. C) The results of a principal components analysis (PCA) decomposing the TD data into two components that collectively explained 90% of variance. D) Scatter plot indicating the lack of a correlation between the 1st Principal Component and the occurrence of TUT$_{CRT}$. E) Scatter plot indicating the significant relationship between the 2nd Principal Component and the occurrence of TUT$_{CRT}$.

kurtosis = 6.80) compared to the PCA component described above (skew = −0.62; kurtosis = 0.67). However, the data do indicate that TUT in low demanding tasks are indicative of less, rather than more, inter temporal discounting demonstrating that the propensity to generate thoughts is related to making more patient long terms choices.

*Structural MRI findings*

a) Cortical thickness substrates of individual differences in TUT. Subjects with higher reports of TUT$_{CRT}$ showed greater cortical thickness in two clusters: a medial cluster encompassed mPFC together with anterior/midcingulate regions and a lateral cluster extending from the inferior and middle frontal regions anteriorly along the operculum to supramarginal regions posteriorly (FWE < 0.05, Fig. 2A). Please note that, after correction for multiple comparisons, findings in medial regions were only significant in the left hemisphere, while similar findings in the right hemisphere did not surpass the stringent threshold set by random field theory. We did not observe any cortical thickness correlates for TUT$_{WM}$ (Fig. 2B). Importantly,

effects for TUT$_{CRT}$ remained robust in the left medial, but not in the lateral prefrontal regions after correction for effects of TUT$_{WM}$ in the same statistical model. In addition, this model controlling for individual differences in TUT$_{WM}$ revealed mPFC and anterior/midcingulate effects in the right hemisphere as well (Fig. 2C). These findings indicate that, after controlling for TUT$_{WM}$, individuals with frequent TUT during the CRT have increased cortical thickness in these anteromedial brain regions relative to those with little TUT during this task. Individuals who tended to titrate their attention to TUT under non-demanding conditions, thus, had greater thickness of mPFC and anterior/midcingulate cortex.

In a separate analysis, we did not observe any significant interaction between gender and TUT$_{CRT}$ on cortical thickness, indicating that our effects were consistent across both males and females.

b) Shared substrates of TUT and TD. One major goal of the present study was to identify shared structural substrates between the capacity to self-generate thought in non-demanding circumstances (i.e., TUT$_{CRT}$) and the ability to make long-term economic decisions (i.e., reduced TD). To this end, we performed three complimentary analyses.
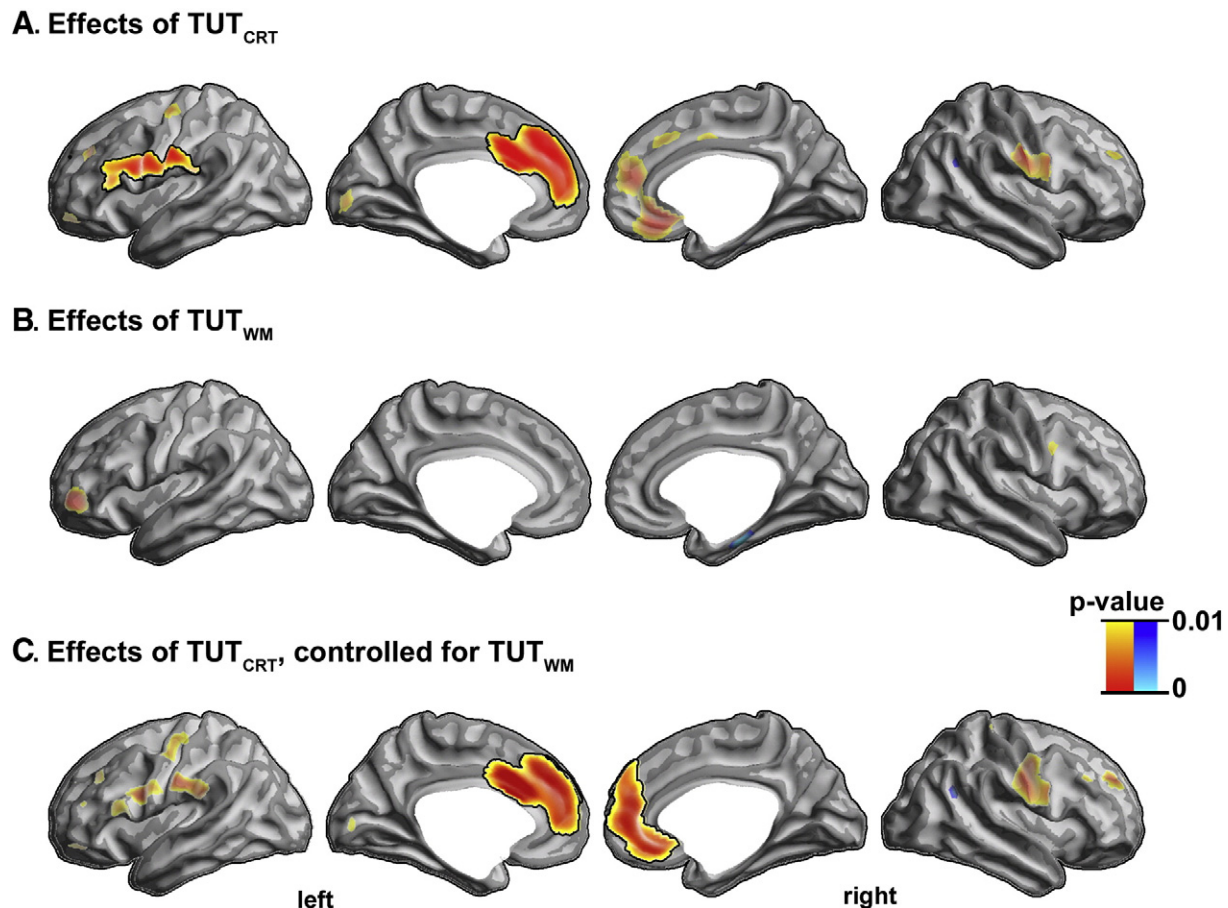
**Fig. 2.** Cortical structural substrates of individual differences in task-unrelated thought (TUT). A) Higher TUT$_{CRT}$ related to increased cortical thickness in medial prefrontal and lateral prefrontal-opercular cortices. B) TUT$_{WM}$ did not show noteworthy cortical thickness correlations. C) Conversely, TUT$_{CRT}$, controlled for effects of TUT$_{WM}$, related to increased cortical thickness in bilateral mPFC and anterior/midcingulate cortices. To correct for multiple comparisons, significances have been thresholded at $p < 0.05$, FWE using random field theory for non-isotropic images (cluster threshold $t > 2.37$, extent threshold = 1.75 resels, black outlines). To illustrate trends, findings at $p < 0.01$, uncorrected (no black outlines, semi-transparent) are also shown.

First, we ran an exploratory overlap analysis between effects of TD, and TUT$_{CRT}$ controlled for TUT$_{WM}$ across the entire cortical surface (each at $p < 0.01$ uncorrected). This analysis yielded, among other regions, overlaps in bilateral mPFC (Fig. 3A). Second, a shared effect of TUT$_{CRT}$ and TD in mPFC was also shown using an independently determined region-of-interest [Talairach space x/y/z = −3/38/13, converted to MNI coordinates using publically available tools (Lancaster et al., 2007)] based on a seminal functional MRI study on TD (Kable and Glimcher, 2007). On the surface, this region was constructed by taking all surface points that were affected by a FWMH = 3 mm surface-based diffusion smoothing applied to the surface-point closest to the coordinates reported by Kable and Glimcher (2007). Thickness of this region was correlated with both TD (left $t = 2.03$, $p < 0.05$, right $t = 1.97$, $p = 0.057$, see Fig. 3B) and TUT$_{CRT}$ controlled for TUT$_{WM}$ (left: $t = 2.61$, $p < 0.02$; right: $t = 2.47$, $p < 0.02$). Third, additional region-of-interest analysis, with regions based on the independently determined and significant clusters of TUT$_{CRT}$ findings (controlled for TUT$_{WM}$; see Fig. 2C), indicated that overall increased thickness in bilateral mPFC and anterior/midcingulate cortices related to individuals' tendencies for less discounting of future rewards during the TD task (left $t = 1.99$, $p < 0.027$; right $t = 2.21$, $p < 0.017$; Fig. 3C). Collectively, these findings suggested that reductions in temporal discounting related to thicker cortices in those medial regions whose thickness also correlated with TUT$_{CRT}$.

## Discussion

Using MRI-based cortical thickness analysis, the current study sought to identify the structural brain substrates underlying individual differences in the capacity to decouple from the here and now and to engage in self-generated thought, as exemplified by TUT. We were particularly interested in mapping substrates of TUT that arise in contexts of relatively low demands on cognitive resources and used a paradigm that allows us to moderate the amount of TUT that participants engage in. In our study, as with many other examples using the same paradigm (Smallwood et al., 2009b, 2011a, 2012b), TUT occurred more frequently during the CRT than in the WM task. Similar results have been observed when investigating the same experience in daily life (Kane et al., 2007), with subjects engaging in more self-generated thought during less demanding situations. Moreover, evidence of a correlation between TUT inside and outside the laboratory (McVay et al., 2009) suggests that it is likely we have captured the tendency to engage in self-generated thought in a context approximating mind-wandering in daily life. We found that self-generated thought processes under low-demanding conditions are associated with individual differences in cortical thickness of mPFC and anterior/midcingulate cortex. These differences in brain structure were also linked to the ability to delay gratification when making monetary decisions. We observed that thickness of a subregion of the anteromedial cluster of significant TUT$_{CRT}$ findings was increased in participants who displayed more patient delay discounting in an economic paradigm, a result that indicates a biological basis for the behavioral correlation we found in this and prior studies (Smallwood et al., 2012b). Our results, therefore, provide novel evidence that structural variations in mPFC and anterior/midcingulate cortex relate to shared individual differences in a capacity to self-generate thoughts decoupled from perceptual input, a process that is thought to be important in allowing individuals to make the predictions necessary to

## A. Whole brain analysis: Effects of TD (blue) and overlap with TUT-CRT (white)
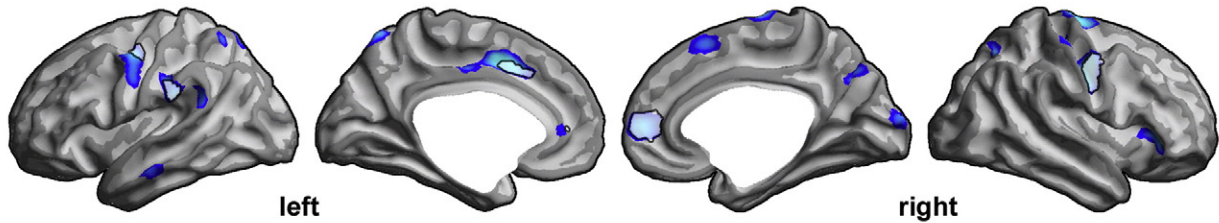


## B. Region-of-interest analysis: Kable and Glimcher (2007)
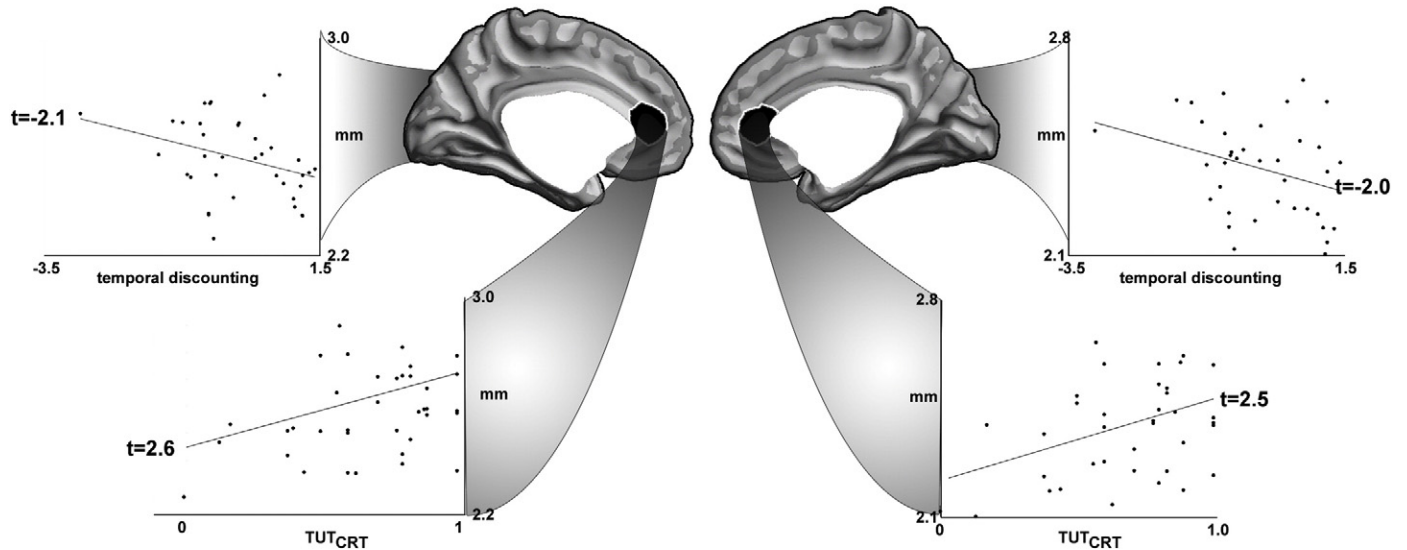


## C. Region-of-interest analysis: effects of TD in clusters of TUT-CRT findings
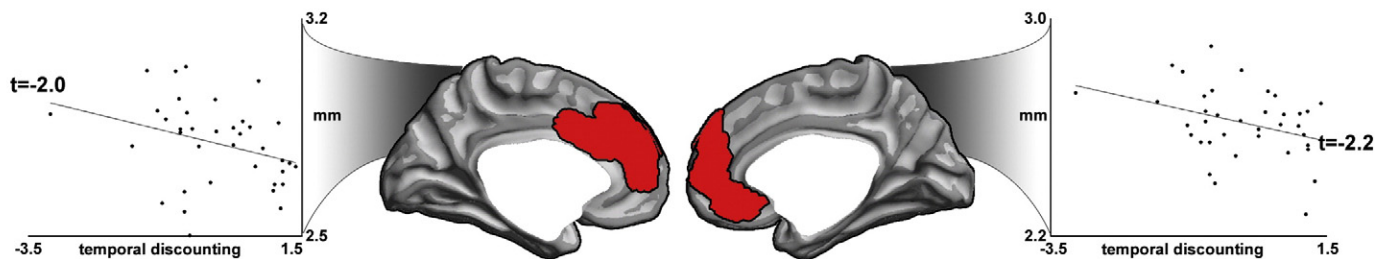


**Fig. 3.** Three separate lines of evidence were suggestive of a shared brain substrate of individual differences in temporal discounting (TD) and task-unrelated thought under low demand conditions (TUT$_{CRT}$). A) Overlap analysis: effects of TD (blue) intersected with effects of TUT$_{CRT}$, controlling for TUT$_{WM}$ (see Fig. 2C), in several regions (white) among them bilateral mPFC. Both maps were thresholded at p < 0.01 uncorrected, prior to forming the overlap. B) Region-of-interest analysis: effects of TD and TUT$_{CRT}$, controlling for TUT$_{WM}$ in an independently determined region-of-interest, based on previously published coordinates (Kable and Glimcher, 2007). C) Region-of-interest analysis: effects of TD on mean thickness in clusters of significant TUT$_{CRT}$ effects (see Fig. 2C).

navigate the complex social environment in which they exist (Amodio and Frith, 2006; Frith and Frith, 1999; Mitchell, 2009; Schacter et al., 2012).

Our structural MRI findings extend previous functional MRI work, which has shown that mPFC and anterior/midcingulate cortex are involved in self-generated TUT (Christoff et al., 2009; Stawarczyk et al., 2011) and in processes by which individuals overcome the tendency to discount a temporally distant reward (Benoit et al., 2011; Kable and Glimcher, 2007). This shared association is consistent with recent component process accounts of both of these phenomena (Amodio and Frith, 2006; Frith and Frith, 1999, 2003; Mitchell, 2009; Peters and Buchel, 2011; Schacter et al., 2012; Smallwood et al., 2012a), which

suggest that TUT and TD share a common dependency on controlled processing of information simulated from memory. This capacity could allow individuals to decouple attention from the here and now and to engage in deliberative processes on information that is not available in the immediate external environment.

The mPFC and also anterior parts of the cingulate cortex have been shown to be involved in a variety of self-generated mental states, including autobiographical memory, mentalizing, and future planning (Amodio and Frith, 2006; Frith and Frith, 1999; Mitchell, 2009; Schacter et al., 2012). A recent study by Benoit and colleagues reported a diminished tendency to discount future rewards in participants who had simulated their future consumption. Using functional MRI, they

demonstrated that mPFC activation correlated with the magnitude of simulated future rewards, which in turn accounted for a significant proportion of the effect of simulation on discounting (Benoit et al., 2011). On the one hand, we suspect that this brain region may be common to both TUT and reduced TD because it is important in the evaluation of information from memory. Anatomically, these anteromedial regions would be well-positioned for this role, given their dense connectivity with medial temporal memory systems, such as the hippocampus and parahippocampal gyrus (Margulies et al., 2007; Petrides and Pandya, 2007; Powell et al., 2004; Squire et al., 2004). On the other hand, mPFC and anterior/midcingulate cortex are also key regions in domain-general conflict monitoring and control processes (Botvinick et al., 1999, 2004; Pardo et al., 1990; Paus, 2001; Petrides and Pandya, 2007; Shackman et al., 2011). Such control may help to suppress habitual responses to external stimuli and facilitate the process of perceptual decoupling by insulating conscious memory-driven processing from ongoing external input (Burgess et al., 2007; Gilbert et al., 2006; Smallwood and Andrews-Hanna, 2013; Smallwood et al., 2012a). Under conditions when attention is directed to an external goal in order to perform an action, processes in cingulate cortex could help focus attention on perceptual input and hence limit disruptions to external task performance that can occur due to self-generated thought. This could account for the association between thickness in these regions and elevations in self-generated TUT that occur in the easy but not more demanding task.

Altogether, the current structural MRI findings add to an emerging body of evidence that a broad class of cognition depends upon the cooperation between control and memory process that allows humans to devote conscious attention to be devoted to thoughts and feelings generated from memory (Peters and Buchel, 2011; Schacter et al., 2012; Smallwood et al., 2012a). Consistent with this emerging picture, mPFC and anterior/midcingulate regions have been shown to be co-activated under conditions that share similar features such as when social information is maintained over a delay (Meyer et al., 2012), when making autobiographical plans (Spreng and Schacter, 2011; Spreng et al., 2010), or when individuals generate a creative solution to a problem (Ellamil et al., 2012).

While providing correlative evidence for a structural substrate for individual differences in the self-generation of thought, the causality underlying this biological link is unclear. Variations in brain structure in healthy individuals have been shown to be driven by genetic factors (Lenroot et al., 2009; Schmitt et al., 2009), developmental processes (Raznahan et al., 2011; Shaw et al., 2006; Steinbeis et al., 2012), and experience-dependent plasticity (Draganski et al., 2006; Engvig et al., 2010); also, there are likely to be several phenotypic characteristics that can account for an individuals' capacity to self-generate thought under non-demanding conditions. For example, people with better cognitive control abilities, as assessed by working memory capacity, tend to engage in TUT under non-demanding conditions (Levinson et al., 2012). This may reflect the important role that control processes play in coordinating self-generated thought, especially when an external task lacks compelling demands (Smallwood and Schooler, 2006; Smallwood et al., 2012a). In this light it is important to note that individual who reported greater TUT in the CRT had greater cortical thickness in a left lateral region of prefrontal cortex that is often associated with the control of information in memory and working memory (Owen et al., 1998; Stern et al., 2000). As overcoming the temptations of an immediate reward depends on neural regions associated with control, as well as those linked to simulation (Schacter et al., 2012), it seems plausible that individual variation in greater cognitive control may mediate the linkage between TUT under nondemanding conditions and reduced temporal discounting.

Finally, several studies have shown that experience-dependent processes can also change cortical structure (Draganski et al., 2006; Engvig et al., 2010). In the case of self-generated thought, such plastic changes could occur because of an individual's habitual tendency to engage in task unrelated thought under conditions when the external environment fails to provide sufficiently compelling environmental input. Irrespective of the cause underlying structural differences, our data suggest that the processes by which self-generated information is prioritized may have its roots in the structure of mPFC and anterior/midcingulate regions. Future studies examining longitudinal changes in the propensity to self-generate TUT, as well as an exploration of their genetic basis, will be important in identifying the precise mechanism that explains the role of these regions in this important form of human cognition.

## Acknowledgments

## References

Addis, D.R., Wong, A.T., Schacter, D.L., 2007. Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. Neuropsychologia 45, 1363–1377.

Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. Nat. Rev. Neurosci. 7, 268–277.

Baird, B., Smallwood, J., Schooler, J.W., 2011. Back to the future: autobiographical planning and the functionality of mind-wandering. Conscious. Cogn. 20, 1604–1611.

Baird, B., Smallwood, J., Mrazek, M.D., Kam, J.W., Franklin, M.S., Schooler, J.W., 2012. Inspired by distraction: mind wandering facilitates creative incubation. Psychol. Sci. 23, 1117–1122.

Baird, B., Smallwood, J., Gorgolewski, C., Margulies, D.S., 2013. Dissociable Networks Support Meta Cognition for Perception and Memory. J. Neurosci. 33 (42), 16657–16665.

Benoit, R.G., Gilbert, S.J., Burgess, P.W., 2011. A neural mechanism mediating the impact of episodic prospection on farsighted decisions. J. Neurosci. 31, 6771–6779.

Bernhardt, B.C., Bernasconi, N., Concha, L., Bernasconi, A., 2010. Cortical thickness analysis in temporal lobe epilepsy: reproducibility and relation to outcome. Neurology 74, 1776–1784.

Bernhardt, B.C., Klimecki, O.M., Leiberg, S., Singer, T., 2013. Structural Covariance Networks of Dorsal Anterior Insula Predict Females' Individual Differences in Empathic Responding. Cereb. Cortex (in press).

Bernhardt, B.C., Valk, S., Silani, G., Bird, G., Frith, U., Singer, T., 2013. Selective Disruption of Socio-Cognitive Structural Brain Networks in Autism and Alexithymia. Cereb. Cortex (in press).

Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., Cohen, J.D., 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature 402, 179–181.

Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci. 8, 539–546.

Buckner, R.L., Carroll, D.C., 2007. Self-projection and the brain. Trends Cogn. Sci. 11, 49–57.

Burgess, P.W., Gilbert, S.J., Dumontheil, I., 2007. Function and localization within rostral prefrontal cortex (area 10). Philos. Trans. R. Soc. Lond. B Biol. Sci. 362, 887–899.

Christoff, K., Gordon, A.M., Smallwood, J., Smith, R., Schooler, J.W., 2009. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. Proc. Natl. Acad. Sci. U. S. A. 106, 8719–8724.

Colom, R., Burgaleta, M., Roman, F.J., Karama, S., Alvarez-Linera, J., Abad, F.J., Martinez, K., Quiroga, M.A., Haier, R.J., 2013. Neuroanatomic overlap between intelligence and cognitive factors: morphometry methods provide support for the key role of the frontal lobes. Neuroimage 72, 143–152.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.

Draganski, B., Gaser, C., Kempermann, G., Kuhn, H.G., Winkler, J., Buchel, C., May, A., 2006. Temporal and spatial dynamics of brain structure changes during extensive learning. J. Neurosci. 26, 6314–6317.

Ellamil, M., Dobson, C., Beeman, M., Christoff, K., 2012. Evaluative and generative modes of thought during the creative process. Neuroimage 59, 1783–1794.

Engvig, A., Fjell, A.M., Westlye, L.T., Moberget, T., Sundseth, O., Larsen, V.A., Walhovd, K.B., 2010. Effects of memory training on cortical thickness in the elderly. Neuroimage 52, 1667–1676.

Fischl, B., Sereno, M., Dale, A., 1999. Cortical surface-based analysis II: inflation, flattening, and surface-based coordinate system. Neuroimage 9, 195–207.

Frith, C.D., Frith, U., 1999. Interacting minds—a biological basis. Science 286, 1692–1695.

Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. Philos. Trans. R. Soc. Lond. B Biol. Sci. 358, 459–473.

Gilbert, S.J., Spengler, S., Simons, J.S., Steele, J.D., Lawrie, S.M., Frith, C.D., Burgess, P.W., 2006. Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. J. Cogn. Neurosci. 18, 932–948.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. Neuroimage 32, 180–194.

Hassabis, D., Maguire, E.A., 2007. Deconstructing episodic memory with construction. Trends Cogn. Sci. 11, 299–306.

Huijbers, W., Pennartz, C.M., Cabeza, R., Daselaar, S.M., 2011. The hippocampus is coupled with the default network during memory retrieval but not during memory encoding. PLoS One 6, e17463.

Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. Nat. Neurosci. 10, 1625–1633.

Kane, M.J., Brown, L.H., McVay, J.C., Silvia, P.J., Myin-Germeys, I., Kwapil, T.R., 2007. For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. Psychol. Sci. 18, 614–621.

Karama, S., Colom, R., Johnson, W., Deary, I.J., Haier, R., Waber, D.P., Lepage, C., Ganjavi, H., Jung, R., Evans, A.C., 2011. Cortical thickness correlates of specific cognitive performance accounted for by the general factor of intelligence in healthy children aged 6 to 18. Neuroimage 55, 1443–1453.

Killingsworth, M.A., Gilbert, D.T., 2010. A wandering mind is an unhappy mind. Science 330, 932.

Kim, J.S., Singh, V., Lee, J.K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J.M., Kim, S.I., Evans, A.C., 2005. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. Neuroimage 27, 210–221.

Kuperberg, G.R., Broome, M.R., McGuire, P.K., David, A.S., Eddy, M., Ozawa, F., Goff, D., West, W.C., Williams, S.C., van der Kouwe, A.J., Salat, D.H., Dale, A.M., Fischl, B., 2003. Regionally localized thinning of the cerebral cortex in schizophrenia. Arch. Gen. Psychiatry 60, 878–888.

Lancaster, J.L., Tordesillas-Gutierrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J.C., Fox, P.T., 2007. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum. Brain Mapp. 28, 1194–1205.

Lenroot, R.K., Schmitt, J.E., Ordaz, S.J., Wallace, G.L., Neale, M.C., Lerch, J.P., Kendler, K.S., Evans, A.C., Giedd, J.N., 2009. Differences in genetic and environmental influences on the human cerebral cortex associated with development during childhood and adolescence. Hum. Brain Mapp. 30, 163–174.

Lerch, J.P., Evans, A.C., 2005. Cortical thickness analysis examined through power analysis and a population simulation. Neuroimage 24, 163–173.

Levinson, D.B., Smallwood, J., Davidson, R.J., 2012. The persistence of thought: evidence for a role of working memory in the maintenance of task-unrelated thinking. Psychol. Sci. 23, 375–380.

MacDonald, D., Kabani, N., Avis, D., Evans, A.C., 2000. Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. Neuroimage 12, 340–356.

Margulies, D.S., Kelly, A.M., Uddin, L.Q., Biswal, B.B., Castellanos, F.X., Milham, M.P., 2007. Mapping the functional connectivity of anterior cingulate cortex. Neuroimage 37, 579–588.

Mason, M.F., Norton, M.I., Van Horn, J.D., Wegner, D.M., Grafton, S.T., Macrae, C.N., 2007. Wandering minds: the default network and stimulus-independent thought. Science 315, 393–395.

McVay, J.C., Kane, M.J., 2011. Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. J. Exp. Psychol. Gen. 141, 302–320.

McVay, J.C., Kane, M.J., Kwapil, T.R., 2009. Tracking the train of thought from the laboratory into everyday life: an experience-sampling study of mind wandering across controlled and ecological contexts. Psychon. Bull. Rev. 16, 857–863.

Meier, S., Sprenger, C.D., 2012. Time discounting predicts creditworthiness. Psychol. Sci. 23, 56–58.

Meyer, M.L., Spunt, R.P., Berkman, E.T., Taylor, S.E., Lieberman, M.D., 2012. Evidence for social working memory from a parametric functional MRI study. Proc. Natl. Acad. Sci. U. S. A. 109, 1883–1888.

Mitchell, J.P., 2009. Inferences about mental states. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 1309–1316.

Owen, A.M., Stern, C.E., Look, R.B., Tracey, I., Rosen, B.R., Petrides, M., 1998. Functional organization of spatial and nonspatial working memory processing within the human lateral frontal cortex. Proc. Natl. Acad. Sci. U. S. A. 95, 7721–7726.

Pardo, J.V., Pardo, P.J., Janer, K.W., Raichle, M.E., 1990. The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. Proc. Natl. Acad. Sci. U. S. A. 87, 256–259.

Paus, T., 2001. Primate anterior cingulate cortex: where motor control, drive and cognition interface. Nat. Rev. Neurosci. 2, 417–424.

Peters, J., Buchel, C., 2011. The neural mechanisms of inter-temporal decision-making: understanding variability. Trends Cogn. Sci. 15, 227–239.

Petrides, M., Pandya, D.N., 2007. Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. J. Neurosci. 27, 11573–11586.

Powell, H.W., Guye, M., Parker, G.J., Symms, M.R., Boulby, P., Koepp, M.J., Barker, G.J., Duncan, J.S., 2004. Noninvasive in vivo demonstration of the connections of the human parahippocampal gyrus. Neuroimage 22, 740–747.

Raznahan, A., Lerch, J.P., Lee, N., Greenstein, D., Wallace, G.L., Stockman, M., Clasen, L., Shaw, P.W., Giedd, J.N., 2011. Patterns of coordinated anatomical change in human cortical development: a longitudinal neuroimaging study of maturational coupling. Neuron 72, 873–884.

Rosas, H.D., Liu, A.K., Hersch, S., Glessner, M., Ferrante, R.J., Salat, D.H., van der Kouwe, A., Jenkins, B.G., Dale, A.M., Fischl, B., 2002. Regional and progressive thinning of the cortical ribbon in Huntington's disease. Neurology 58, 695–701.

Schacter, D.L., Addis, D.R., Buckner, R.L., 2008. Episodic simulation of future events: concepts, data, and applications. Ann. N. Y. Acad. Sci. 1124, 39–60.

Schacter, D.L., Addis, D.R., Hassabis, D., Martin, V.C., Spreng, R.N., Szpunar, K.K., 2012. The future of memory: remembering, imagining, and the brain. Neuron 76, 677–694.

Schilling, C., Kuhn, S., Paus, T., Romanowski, A., Banaschewski, T., Barbot, A., Barker, G.J., Bruhl, R., Buchel, C., Conrod, P.J., Dalley, J.W., Flor, H., Ittermann, B., Ivanov, N., Mann, K., Martinot, J.L., Nees, F., Rietschel, M., Robbins, T.W., Smolka, M.N., Strohle, A., Kathmann, N., Garavan, H., Heinz, A., Schumann, G., Gallinat, J., 2013. Cortical thickness of superior frontal cortex predicts impulsiveness and perceptual reasoning in adolescence. Mol. Psychiatry 18, 624–630.

Schmitt, J.E., Lenroot, R.K., Ordaz, S.E., Wallace, G.L., Lerch, J.P., Evans, A.C., Prom, E.C., Kendler, K.S., Neale, M.C., Giedd, J.N., 2009. Variance decomposition of MRI-based covariance maps using genetically informative samples and structural equation modeling. Neuroimage 47, 56–64.

Schooler, J.W., Smallwood, J., Christoff, K., Handy, T.C., Reichle, E.D., Sayette, M.A., 2011. Meta-awareness, perceptual decoupling and the wandering mind. Trends Cogn. Sci. 15, 319–326.

Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., Davidson, R.J., 2011. The integration of negative affect, pain and cognitive control in the cingulate cortex. Nat. Rev. Neurosci. 12, 154–167.

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., Evans, A., Rapoport, J., Giedd, J., 2006. Intellectual ability and cortical development in children and adolescents. Nature 440, 676–679.

Smallwood, J., 2013. Distinguishing how from why the mind wanders: a process-occurrence framework for self-generated mental activity. Psychol. Bull. 139, 519–535.

Smallwood, J., Andrews-Hanna, J.R., 2013. Not All Minds that Wander are Lost: The Importance of a Balanced Perspective on the Mind-Wandering State towards a Balanced Perspective of the Mind-Wandering State. Front. Psychol. 4, 441.

Smallwood, J., Schooler, J.W., 2006. The restless mind. Psychol. Bull. 132, 946–958.

Smallwood, J., Fitzgerald, A., Miles, L.K., Phillips, L.H., 2009a. Shifting moods, wandering minds: negative moods lead the mind to wander. Emotion 9, 271–276.

Smallwood, J., Nind, L., O'Connor, R.C., 2009b. When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. Conscious. Cogn. 18, 118–125.

Smallwood, J., Brown, K.S., Tipper, C., Giesbrecht, B., Franklin, M.S., Mrazek, M.D., Carlson, J.M., Schooler, J.W., 2011a. Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. PLoS One 6, e18298.

Smallwood, J., Schooler, J.W., Turk, D.J., Cunningham, S.J., Burns, P., Macrae, C.N., 2011b. Self-reflection and the temporal focus of the wandering mind. Conscious. Cogn. 20, 1120–1126.

Smallwood, J., Brown, K., Baird, B., Schooler, J.W., 2012a. Cooperation between the default mode network and the frontal–parietal network in the production of an internal train of thought. Brain Res. 1428, 60–70.

Smallwood, J., Ruby, F.J., Singer, T., 2012b. Letting go of the present: mind-wandering is associated with reduced delay discounting. Conscious. Cogn. 22, 1–7.

Spreng, R.N., Schacter, D.L., 2011. Default network modulation and large-scale network interactivity in healthy young and old adults. Cereb. Cortex 22, 2610–2621.

Spreng, R.N., Stevens, W.D., Chamberlain, J.P., Gilmore, A.W., Schacter, D.L., 2010. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. Neuroimage 53, 303–317.

Squire, L.R., Stark, C.E., Clark, R.E., 2004. The medial temporal lobe. Annu. Rev. Neurosci. 27, 279–306.

Stawarczyk, D., Majerus, S., Maquet, P., D'Argembeau, A., 2011. Neural correlates of ongoing conscious experience: both task-unrelatedness and stimulus-independence are related to default network activity. PLoS One 6, e16997.

Steinbeis, N., Bernhardt, B.C., Singer, T., 2012. Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. Neuron 73, 1040–1051.

Stern, C.E., Owen, A.M., Tracey, I., Look, R.B., Rosen, B.R., Petrides, M., 2000. Activity in ventrolateral and mid-dorsolateral prefrontal cortex during nonspatial visual working memory processing: evidence from functional magnetic resonance imaging. Neuroimage 11, 392–399.

Thompson, P.M., Hayashi, K.M., Sowell, E.R., Gogtay, N., Giedd, J.N., Rapoport, J.L., de Zubicaray, G.I., Janke, A.L., Rose, S.E., Semple, J., Doddrell, D.M., Wang, Y., van Erp, T.G., Cannon, T.D., Toga, A.W., 2004. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. Neuroimage 23 (Suppl. 1), S2–S18.

Worsley, K., Andermann, M., Koulis, T., MacDonald, D., Evans, A., 1999. Detecting changes in nonisotropic images. Hum. Brain Mapp. 8, 98–101.

Worsley, K.J., Taylor, J.E., Carbonell, F., Chung, M.K., Duerden, E., Bernhardt, B.C., Lyttelton, O., Boucher, M., Evans, A.C., 2009. SurfStat: a Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory. Neuroimage 47.