When the brain takes a break:
A model-based analysis of mind wandering

Matthias Mittner (geb. Ihrke)
Cognitive Science Center Amsterdam

July, 31st, 2014
ICON 2014, Brisbane
Mind-Wandering is an ubiquitous phenomenon...

- iPhone-app: people are continuously queried about what they are doing
- Frequency of mind-wandering: 40-50% independent of current activity

→ If we spend half our waking time daydreaming, can we assume that our experimental subjects are task-centered at all times?

*Killingsworth & Gilbert, 2010, Science*
What is Mind-Wandering?

Different experimental contexts

1. task-unrelated thoughts (TUT)
2. attentional lapses (failure to perceive/respond)
3. stimulus-independent thoughts (SIT)
4. tuning out vs. zoning out (with and without meta-awareness)
5. ...
Experimental Findings

Mind-wandering

1. decreases with growing task-difficulty
2. increases with growing practice on task (automatization)
3. increases with current concerns (baseline thought-production)
4. increases with alcohol consumption
5. increases with nicotine craving
6. increases with fatigue
7. decreases with working-memory capacity
8. is increased in ADHD-patients
9. is increased in mild depression patients
10. is decreased in older adults
Why and How?

Why? (functions of Mind Wandering)
- future planning (internal practice)
- creativity
- attentional cycling (inherent tendency to shift attention)
- dishabituation (mind wandering as break from current task)

How? (what mechanism is involved)
- executive control is involved (failure vs. resource)
- what kind of control?
Experimental Setup

Stop-Signal Task

- allows to distinguish between different aspects of executive control ("goal-monitoring" and "stopping")
- left/right arrows, response left/right
- beep indicates stop the current response
- stop-signal delay (SSD) adjusted to produce 50% errors
- measures: fMRI, pupil, behaviour
Experimental Setup

Thought-probes

- randomly presented during the course of the experiment (ca. 1 per minute)
- 5-point Likert-scale
- common operationalization of mind-wandering in attention experiments
Goals of this project

Outline

- identify mind-wandering on a single-trial level
- analyse the neural and behavioural signature of Mind-Wandering
  → identify which cognitive processes are impaired using cognitive models of behaviour
Theory: fMRI and Mind-Wandering

potential fMRI correlates

- Default-Mode Network (DMN) and Anticorrelated-Network (ACN)
  - DMN activity increased prior to mind-wandering (Christoff et al., 2009, PNAS)
- DMN/ACN dynamic functional connectivity (dFC) related to vigilance (Thompson et al., 2013, HBM)

Fox et al. (2005), PNAS
Theory: Pupil Data

Potential Correlates

- pupil diameter possibly correlated with locus coeruleus-norepinephrine (LC-NE) activity
- Adaptive Gain Theory (AGT, Aston-Jones et al., 2005)
  - tonic LC-activity: baseline pupil diameter
  - phasic LC-responses: pupil-response function

(Rajkowski et al., 1993) (Aston-Jones & Cohen, 2005)
Summary: Classification

Trial-wise features for Classification

- baseline PD
- pupil response
- pre-stimulus mean
- stimulus-induced transient

ROIs

- ACN (7 ROIs)
- DMN (5 ROIs)

Dynamic FC

- correlation within ACN (21 correlations)
- correlation within DMN (10 correlations)
- correlation bw DMN/ACN (35 correlations)

Classifier

Support Vector Machine

Thought-probes

What did you think about, during the last trial?

on-task

off-task

targets

Matthias Mittner (geb. Ihrke) 31.7.2014
Results: Classification

Classifier optimization

- Support-vector machine with gaussian radial basis functions
- Optimize RBF-SVM parameters \((C, \gamma)\) using AUC criterion
- Recursive feature elimination
- Cross-Subject Crossvalidation Accuracy: 79.5%

⇒ **single-trial** probability of mind wandering for each trial
Results: Classification
Feature Importance/Activation

- **A**
  - Feature activation
  - on-task
  - off-task

- **B**
  - Perturbation score

- **ACNDMNDMN DMN ACN ACNDMN ACN pupil**

Matthias Mittner (geb. Ihrke)
When the brain takes a break
31.7.2014
Results: Classification
Feature Activation

Classifier
Support Vector Machine

on-task trials

off-task trials

Feature activation on vs. off-task

- DMN activity predicts off-task, ACN predicts on-task
- absolute connectivity during off-task stronger
- PD baseline and response reduced during off-task
Cognitive Model

Independent Race Drift-Diffusion Model

- allows to decompose reaction times into cognitive processes (parameters):
  - efficiency of go/stop processes (drift rates, $V, v$)
  - caution (boundary separation, $b$)
  - duration of perception/motor (nondecision time, $t_{er}$)
- goal-monitoring vs. inhibitory processes
Cognitive Model

Method: Bayesian hierarchical Modeling

When the brain takes a break
31.7.2014
Mind-Wandering is

- reflected in a combination of decreased drift-rates and decreased boundary (sign. on posterior modes) → goal-monitoring affected
- more “impulsive” behaviour:
  - behaviour more variable (longer distribution tails)
  - more errors
- inhibitory processes not affected
Conclusion

- Mind wandering can be predicted on the single-trial level (80% accuracy)
- ... using theoretically meaningful, neural variables
- the classification signature agrees with predominant view of DMN influence on MW
- MW affects executive goal-monitoring but not inhibitory processes
Thank you for abstaining from mind wandering!
Preprocessing: functional connectivity

Residual General Linear Model

- voxel activity \( y_i = \beta_1 x_1 + \cdots + \beta_m x_m + \epsilon \)
  incl. task, motion, blinkrate, white-matter and CSF
- subtract estimate from data to obtain residuals
  \( \rho_i = y_i - \sum_i \beta_i x_i \)

ROI definition

- global correlation map with PCC seed
- per-subject definition of ROIs
Method: fMRI (ROI activity and functional connectivity)

Activity before Thought-Probes

→ effects in resting-state activity up to \(\approx 20s\) back
→ use integrated activity over that window

Dynamic functional connectivity

- sliding-window correlation \(\text{corr}_w(\rho_i(t), \rho_j(t))\) for \(t \in W_k = \{k, \ldots, k + w\}\)
- Problem: what is the “correct” window size \(w\)?

→ use 40s window (Shirer et al., 2012, CB)
Method: Pupil Diameter

Baseline Pupil-Diameter
- mean PD [1000, 0] ms before trial onset

Pupil Response
extract pupillary response via GLM

stimulus onsets

IRF \( h(t) = t^{10.1} e^{-\frac{10.1t}{930}} \)

regressors

Irregularities in pupil diameter are analyzed using a Support Vector Machine (SVM) classifier. The baseline pupil diameter is calculated as the mean pupil diameter from 1000 ms before the trial onset until the onset of the stimulus. The pupillary response is extracted using a Generalized Linear Model (GLM) to model the time course of pupil dilation and constriction in response to the stimulus. The impulse response function (IRF) is given by the equation above, where \( h(t) \) represents the pupil response at time \( t \).
Cognitive Model
Independent Race Drift-Diffusion Model

Go-Trial

+ 
500 ms
until response
ITI

Stop-Trial

+ 
SSD

500 ms
until response
ITI

$V_b$
$t_{er}$

$V_a$

$t_{er}$

b
SVM predicts state sequence for each trial $i$:
$$\hat{S} = (\hat{S}_1, \ldots, \hat{S}_N), \hat{S}_i \in \{\text{on, off}\}$$

SVM classification not perfect but: probability for correct prediction
$$P(\hat{S}_i = S_i) := p_{acc}(i)$$

→ model uncertainty as a mixture:
$$p(x_i|\theta) = p_{acc}(i)f(x_i|\hat{S}_i) + (1 - p_{acc}(i))f(x_i|\neg\hat{S}_i)$$
Analysis of frequency of mind wandering

- TUTs increase with time
- TUTs are mainly short (a few trials)
- Good correspondence to previous work (Bastian et al., 2013)
cluster-analysis reveals 3 clusters of subjects with distinct behavioural patterns

1. cluster 1 (N=2): inverse goal-monitoring effect
2. cluster 2 (N=8): no effect
3. cluster 3 (N=10): goal-monitoring effect

different “kinds” of mind wandering involved?
Cognitive Model
Dealing with Classifier Uncertainty

- SVM predicts state sequence for each trial $i$:
  $\hat{S} = (\hat{S}_1, \ldots, \hat{S}_N), \hat{S}_i \in \{\text{on, off}\}$

- SVM classification not perfect but: probability for correct prediction
  $P(\hat{S}_i = S_i) := p_{\text{acc}}(i)$

→ model uncertainty as a mixture:

$$p(x_i|\theta) = p_{\text{acc}}(i)f(x_i|\hat{S}_i) + (1 - p_{\text{acc}}(i))f(x_i|\neg\hat{S}_i)$$