A computational model of focused attention meditation and its transfer to a sustained attention task

Amir J. Moye (amir.moye@psy.unibe.ch)
Institute of Cognitive Psychology, Perception and Research Methods, University of Bern
Fabrikstrasse 8, 3012 Bern, Switzerland

Marieke K. van Vugt (m.k.van.vugt@rug.nl)
Institute of Artificial Intelligence and Cognitive Engineering, University of Groningen
Nijenborgh 9, 9747 AG Groningen, The Netherlands

Abstract

Although meditation and mindfulness practices are widely discussed in the scientific literature, there is little formal theory about the cognitive mechanisms that comprise it. Here we begin to develop such a theory by creating a computational cognitive model of a particular type of meditation: focused attention meditation. This model was created within *Prims*, a cognitive architecture similar to and based on *ACT-R*, which enables us to make predictions about the cognitive tasks that meditation experience may affect. We implemented a model based on an extensive literature review of how the meditation experience unfolds over time. We then subjected the Prims model to a session of the *Sustained Reaction to Response Task*, a task typically used to study sustained attention, a faculty that may be trained with meditation practice. Analyses revealed that the model was significantly more sensitive to detecting targets and non-targets after the meditation practice than before. These results agree qualitatively with empirical findings of a longitudinal study conducted in 2010. These results suggest that our approach to modeling meditation and its effects of cognition is feasible.

***Keywords*:** Focused attention meditation, mindfulness, sustained attention, SART, PRIMS, transfer

Introduction

Meditation consists of a set of mental exercises that have been developed and practiced reaching as far back as 4000 years (Riley, 2004). In the last 50 years there has been more and more interest in the effects of the various meditation styles on cognition and emotion. The spectrum of empirically examined effects has grown quite vast, with some being reasonably well-replicated and of medium to large effects while others have been inconsistent (Khoury, Sharma, Rush, & Fournier, 2015; Sedlmeier et al., 2012). However, there are no comprehensive computational frameworks of meditation and its effect on cognition (e.g., Vago & Silbersweig, 2012).

Meditation is often conceptualized as a family of attentional and emotional regulation exercises, the former being the aspect that virtually all styles share to some degree. However, it needs to be stressed that meditation techniques differ strongly. They originate from distinctive cultures and religions (Buddhism, Hinduism, Taoism, Sufism, Christian Centering Prayer, etc.) as well as secular settings (acceptance and commitment therapy, mindfulness-based stress reduction, mindfulness-based cognitive therapy; Hayes, 2004; Kabat-Zinn, 1990; Teasdale et al., 2000). They can differ greatly concerning the emphasis of the mental faculties used (attention, feeling, reasoning, visualization, etc.), the objects they are focused on (thoughts, images, concepts, internal energy, breath, love, God, etc.; Shear, 2006) and lastly with what aim they are employed (relaxation, heightened sense of well-being; attentional balance, insight, etc.; Lutz, Slagter, Dunne, & Davidson, 2008; Wallace, 1999). That being said, the common typology to categorize this vast family of practices is based on what meditators are purportedly doing from a first-person perspective: 'Focused Attention' (FA) meditation and 'Open Monitoring' (OM) meditation (Lutz et al., 2008). In OM practices – in contrast to FA meditation – there is no clear focus of attention and the task is to be continuously aware of phenomena appearing and to return to this monitoring when one gets caught up with the content.

In this paper, we begin to develop a computational theory of meditation practices by creating a cognitive model of focused attention (FA) meditation, as this kind of meditation is most amenable to computational modeling. In this practice, the meditator brings her/his attention to an object such as the breath, and then monitors with non-judgmental attention whether attention is still there. As soon as the meditator realizes attention has wandered, s/he brings the attention back to the object of focus, minimizing any further mental elaboration.

The particular type of FA meditation that was practiced by the subjects relevant for this article was so-called *Samatha* meditation (MacLean et al., 2010). According to Wallace (1999), the meditation instructor of the retreat, the main goal of this practice is to cultivate a stability and vividness concerning attention. In order to pursue this cultivation there are two crucial faculties that must be refined in turn: *mindfulness* and *introspection*, mindfulness being the primary faculty. In the setting of Samatha, mindfulness may be reduced to the aspects of recollection and steadiness: the ability to remember to sustain the attention on a given object and to remember to return when there has been a distraction nevertheless (Wallace, 1999). Introspection, on the other hand, is the faculty to monitor the meditation process, a type of meta-cognition that is tuned to the detection of increases in phenomenological excitation or laxity. When these two faculties fail, mind wandering may take over: an unintended shift of focus to a sensory or mental event, which then leads to habitual affective responding, which in turn triggers related mental events such as episodic or procedural memories, that then lead to more habitual affective responses and so on (Vago & Silbersweig, 2012).

The meditation model was constrained in two ways: (i) qualitatively through taking testimonials and existing theories on meditation into account and (ii) quantitatively by taking existing data into account. Because meditation itself produces virtually no behavioral output to which one could compare a model output, our model was constrained indirectly by having it predict transfer to a similar task that does produce output. This transfer was compared to empirical data of a three month FA meditation retreat (MacLean et al., 2010). The specific transfer was from multiple FA meditation sessions to a *Sustained Attention to Response Task* (SART; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). The similarity between the modeled and actual transfer effect is then an indirect measure for the fit of the meditation model to the actual meditation process. The rationale here is that an adequate model of meditation would be expected to make reasonably good predictions about transfer to other tasks.

The SART is a useful task to examine the effects of meditation practice, because both meditation and this task involve maintaining attention over a long period. In the SART, typically performance is quite good at first but quickly decreases. This vigilance decrement is characterized by a reduction in speed and accuracy as well as reductions in perceptual sensitivity and increases in response bias (Warm, 1980). According to Lutz et al. (2008) there are significant parallels between conceptualizations of sustained attention in cognitive sciences and processes involved in FA meditation. Moreover, there is consensus between Western scientists and Buddhist scholars that both processes require "skills involved in monitoring the focus of attention and detecting distraction, disengaging attention from the source of distraction, and (re)directing and engaging attention to the intended object" (Lutz et al., 2008, p. 2).

Computational models for the SART already exist (Gunzelmann, Gross, Gluck, & Dinges, 2009; van Vugt, Taatgen, Sackur, & Bastian, 2015). The SART model created for this paper was inspired by the model by van Vugt et al. (2015), which–contrary to other models that leave mind-wandering abstract–models mind-wandering explicitly as a cognitive process of memory retrieval. The advantage of modeling mind-wandering explicitly is that it allows you to model the actual thoughts that are mind-wandered about, and the change in attitude towards these thoughts that is so characteristic of meditation practice (Desbordes et al., 2015; Vago & Silbersweig, 2012). Even though there are several comprehensive theoretical frameworks of meditation (e.g., Vago & Silbersweig, 2012), to the best of our knowledge there are not yet any computational models of meditation, let alone FA meditation.

We implemented our model in the *Prims architecture* (Primitive Information Processing Elements; Taatgen, 2013). It is a recent extension of the well-established *Adaptive Control of Thought – Rational*, or *ACT-R* (Adaptive Control of Thought-Rational; Anderson & Lebiere, 2012) and has been developed to be able to explain transfer between different cognitive tasks, which is crucial for our project. As in ACT-R, cognitive processing is distributed across specialized modules, which are implied by some theories of cognition (Anderson & Lebiere, 2012):

- A goal module, which stores active goals and applies their influence.
- An input module, which models perception (e.g., vision)
- An output module, which model outward actions (e.g., button presses)
- A retrieval module, which models declarative memory and memory retrieval processes.
- A working memory module, which stores information that is immediately accessible and intermediate steps in calculations

Cognitive processing itself takes place in cycles of applying if-then-rules. These rules are called *operators* in Prims (and *productions* in ACT-R). In every cycle, the information in the buffers of the modules is compared to the conditions of the operators. If multiple operators have conditions that fit the information in the system, a competition between them occurs and the operator with the highest activity – which depends among other factors on a baseline activity plus a random noise variable – is chosen to be executed.

Method

When the model is run for several rounds it simulates roughly four processes that a meditator cycles through:

1. Remembering (or keeping in mind) what is supposed to be done again and again: In this case, this is the task of being aware of the breath.
2. Being aware of breath sensations, which is simulated as copying the perception into working memory.
3. Remembering something else and wandering off into daydreams, worries, etc.
4. Remembering to come back to the task when one has wandered off.

The model does this by assuming two competing goals[1] – focusing on the breath (the focus goal) and mind wandering (the wander goal) – which each have operators associated with them (van Vugt et al., 2015). Which operator wins depends on three factors in this model: the baseline activation of the operator, the random activation added and the spreading activation from the goal it is associated with. Goals can furthermore be activated or deactivated by operator actions (a unique feature of Prims that ACT-R does

[1] These goals – especially the goal to mind wander – are not necessarily explicit/conscious to the individual.

not have). In the latter case, their activation is automatically 0. This does not mean that operators associated with an inactive goal cannot win a competition; it just makes it a lot less likely.

As can be seen in Figure 1, all of the operators are triggered by the retrieval of the last cycle and as can be seen in Table 1, there are three kinds of memory chunks. 30 of them are meant to model mind-wandering contents (not necessarily single memories but rather representative instances of narratives or overarching themes). The 31st is the memory of the meta-task, which is the memory of refreshing the goal itself before checking what the low-level task at hand is. The 32nd is the memory of the low-level task, which entails feeling the breath.

Table 1: The three types of memories in the declarative memory of the meditation model and their slots.

Meta-task (n=1)	Task (n=1)	Mind-wandering (n=30)
Memory	Memory	Memory
Intention	Intention	Mind-wandering
Meta-task	Task	Memory-4*
Focus	Breath	Approach*

Note: * These are examples. The memory slot ranges from 'Memory-1' to 'Memory-30' and the valence slot can contain 'Approach', 'Avoid' or 'Stay'.

A mind-wandering memory could have the following slot contents: Memory, Mind-wandering, Memory-17, Avoid. The first slot indicates that this chunk is a memory, which is a very general label to allow for general requests. The second slot distinguishes the mind-wandering chunks from the memories of intentions, while the third slot is a placeholder for a specific memory topic (e.g. 'Memory-21' might be a future-oriented and attractive topic – going on vacation). Finally, the fourth slot contains the valence or motivational connotation. Both intention memories have lower activations to begin with, 1.00 as opposed to the mind-wandering chunk's average activation of 3.07. This models the intention memories being less salient and engaging (at first) than the mind-wandering memories.

The model starts off with the focus goal activated and 'Breath' in the input buffer (which remains there). As nothing has been retrieved, the retrieval operators of both goals will compete. At this point the focus operator will usually win, as the wander goal is not active yet. If it does, it requests a general memory and since it has associations with the task and meta-task memories, they have a better chance than the daydream memories of being recalled (if they have the same baseline activation anyway). If the task memory is remembered this directly triggers being aware of the breath, however if the meta-task memory is recalled this first triggers the refresh-focus-operator. This activates the focus goal if it was inactive or reinforcing it if it was already active. Next the opposing goal is deactivated if it is active and the concrete task at hand is requested, modeling a meta-cognitive process that consist of reinforcing the goal to focus and remembering the task to focus on. After feeling the breath nothing is retrieved and the retrieval operators once again are triggered. If the wander operator wins it will initiate a similar process as outlined for the focus goal, thereby reinforcing the wander goal. Once a goal has been activated its operators tends to go into a stable loop. However, as can be seen in figure 1 there are multiple interception points to interrupt this.

The model of sustained attention simulates the

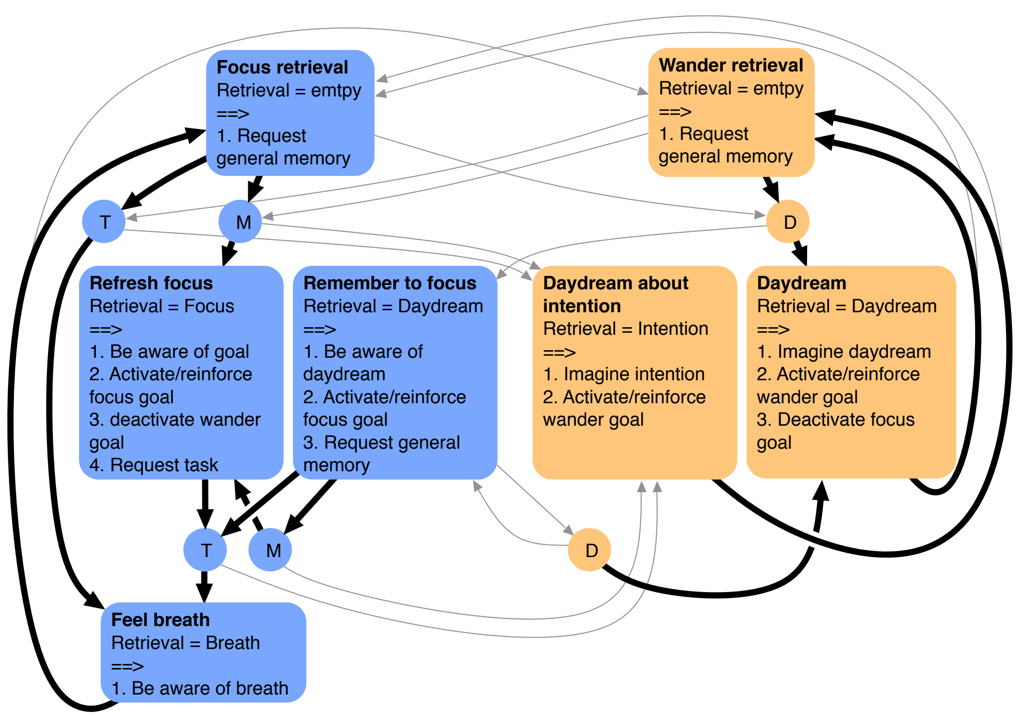

Figure 1: Meditation model. Blue objects are related to the focus goal, yellow ones are related to the wander goal. The boxes are operators, while the small circles are memories that are retrieved due to a request by an operator. 'T' stands for task, 'M' stands for meta-task, 'D' stands for daydream. The arrows represent possible transitions. Thick black arrows represent high probability, while thin gray arrows indicate lower probability. The represented probabilities always signify the chances if both goals were active and the memories had similar baseline activations.

performance of the meditators in a SART that the participants of the meditation retreat performed (MacLean et al., 2010). It consists of frequent non-targets (long lines, with 90% probability) and rare targets (short lines, with 10% probability). The screen switched between the display of a mask (1.55-2.15s) and the display of a stimulus (0.15s). There was a practice block of 120 trials and 4 contiguous test blocks of 120 trials each, which lasted for about 18min. The main measure was *A*' (Stanislaw & Todorov, 1999), a measure of sensitivity combining hit rates and false alarms.

The model (Figure 2) is made up of operators for modeling the mind wandering as well as operators for modeling the execution of the SART task. The operators for mind wandering are almost identical to their respective copies from the meditation model. In a sense, the model consists of SART operators (identifying the stimulus, pressing, etc.) and a modification of the meditation model missing the primary and secondary focus operators.

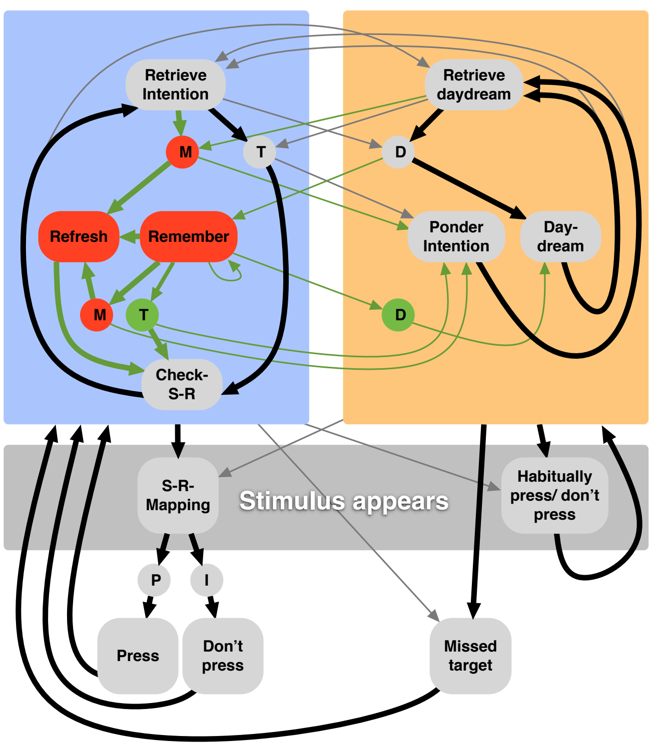

Figure 2: Model of the SART after transfer. 'T' stands for task, 'M' stands for meta-task and 'D' stands for daydream. The red objects are the transferred operators and the meta-task memory. The green memories and transition lines appear in the diagram as a consequence of this transfer.

The transfer consisted of copying the meta-task memory and two meditation model operators into the SART model, transferring the following processes: Reinforcing/activating the focus goal, deactivating the wander goal, reinforcing the focus related memories and the process of remembering the task at hand when mind wandering. Importantly, the low-level task and its memory differed from their counterparts in the meditation model: in the SART the low-level task was to check the stimulus-response-mapping in case a stimulus appeared.

Results

Prims has a vast spectrum of parameters, most of which influence the performance of the models. A majority of them were kept at the default level, while some were adjusted to allow for both models to perform at least somewhat realistically. Specifically, the activation noise was set to 0.4 (default is 0.1), which allowed for slower transitions, more interference and shorter loops. The amount of goal buffer spreading activation was set to 0.75 (default is 1), which decreases the impact the goal activation/deactivation has, with similar effects as the increased activation noise parameter. The amount of working memory buffer spreading activation was set to 0.3 (default is 0), which allows for association between daydreams during mind wandering. The latency factor was set to 0.15 (default is 0.2) to make the SART model faster in responding to the stimulus. The learning parameter for production compilation was set to 0.2 (default 0.1) to allow the SART model to assemble the prims faster in the training phase.

The meditation model was tested for a simulated 18 hours at which point it seemed to have reached a dynamic equilibrium (representing the process of learning to stay focused on the breath). The analyses reported here pertain to only one run, as there was very little variation between the runs. As can be seen in figure 3, the model starts off with a lot of mind wandering but slowly begins to shift to more focus and then drops below the rising focus percentage out at about 5 hours. In the end almost all retrieved memories are focus related.

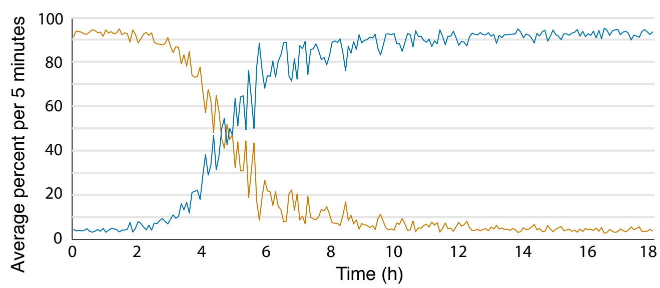

Figure 3: The average percentage (of 5-minute periods) of focus (blue) and wander (yellow) operators during a simulated 18h run.

The SART model was run for 1 training block and 4 test blocks like in the empirical study. The results presented are the average of 30 runs, as the SART model was somewhat variable in its performance, partly due to the relatively short simulated time span (18 min as opposed to 18 hours for the meditation model).

The SART model with transfer was run with a meta-task memory at the low starting activation level of the meditation model: 1.00. As can be seen in Figure 4, the mind wandering percentage is lower, while the focus percentage

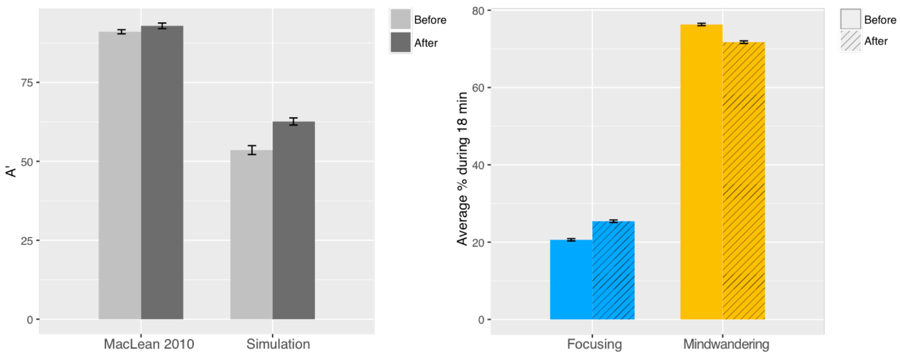

Figure 4: Measures in the SART before and after meditation training. The plot on the left compares the results found in the study by MacLean et al. (2010) before and after the retreat with the performance of the model before and after transfer of two operators and a memory at an activation level of 1.00. The gray bars represent the average sensitivity *A'*. The plot on the right compares the average percentage of focus and mind wandering operators respectively during 18min runs.

has increased. Furthermore, the hit rate was increased and the false alarm rate was lower (not displayed in the figure), leading to an increased sensitivity. An independent t-test of the mean sensitivity over time revealed that the difference was highly significant ($t(58) = 4.49$, $p < 0.001$) and that Cohen's effect size of the difference was large ($d = 1.18$). Examination of the Q-Q plot and the Shapiro-Wilk test showed no significant deviation from normality ($W(60) = 0.99$, p = 0.85). The difference was even more pronounced when the meta-task memory was transferred at above average activation levels (4.50): $t(47.82) = 14.05$, $p < 0.001$, $d = 3.69$. The assumption of normality was rejected ($W(60) = 0.92$, $p = 0.001$). Therefore a bootstrap test was conducted, which corroborated the significance of the effect ($p = 0.001$).

Discussion

This paper set out to explore the processes underlying FA meditation by creating a cognitive model to simulate it. To constrain the model and test its plausibility, a cognitive model of a SART was analyzed before and after the transfer of two meditation operators and an affiliated memory.

The meditation model transitions from mainly mind wandering to being almost entirely focused on the task at hand. There seem to be two main causes for this development: The increasing dominance of the meta-task memory over the task memory as well as the increasing dominance of both intention memories over the mind wandering memories. The fact that the meta-task memory becomes stronger than the task memory leads to more instances of the following sequence: focus retrieval → refresh focus → feel breath, and less of this sequence: focus retrieval → feel breath. This in turn allows for more reinforcement of the focus goal and the meta-task memory because the refresh-focus-operator involves goal management actions and imagination (strengthens the memory). The second cause–the domination of the intention memories over the mind wandering memories–leads to more of their retrieval and less retrieval of the mind wandering memories. In other words, it decreases the probability of interference by mind wandering memories and increases the probability of the intention memories (mostly the meta-task memory) interfering with the wander-retrieval.

This raises the question why the meta-task and the task memory increase in activation so dramatically over time. The intention memories probably increased because they are retrieved a lot more than any single mind-wandering memory. Even though the mind-wandering memories as a whole are retrieved a lot more frequently at first than the intentions and even though they spread the resulting reinforcement amongst each other to some degree (due to their associations), the reinforcement per single mind-wandering memory is a lot smaller than for the meta-task. What gives the mind-wandering memories the upper hand at first–their numbers–becomes a handicap as the reinforcement they receive is spread too evenly among them. This has interesting implications. It could mean that an important aspect of how FA meditation calms the mind lies in its simplicity and unidirectionality: it only focuses on a small group of memories, while mind-wandering has a broad focus. It could indicate that if the goal management strategy is such that it is sufficient for combating mind-wandering loops and interference–even if only rarely at first–it can reinforce its associated memories, causing it to be more effective in turn, which leads to more reinforcement and so on. In other words, if the goal management strategy is effective enough in the beginning (even if only barely) it can create a feedback loop. And while the mind-wandering process creates a feedback loop as well, it is less effective, presumably because the loop is a lot more dispersed.

What is interesting about mind-wandering is that it seems to creep up stealthily and is often easy to snap out of, but only for a few moments, which reflects what we think are two core factors in mind-wandering's longevity: tenacity and momentum. The meditation model explored in this paper suggests that FA meditation functions on the same principles supplemented with the benefits of unidirectionality. Yet, what this model leaves out is that mind-wandering is typically not a deliberate choice, while a main aspect of FA meditation is the conscious, voluntary and therefore effortful deciding from moment to moment. The model cannot distinguish between bringing something to mind consciously and something appearing on its own (Seli, Carriere, & Smilek, 2015).

Possibly the central question is how plausible the meditation model is. The meditation model was almost entirely constrained by internal consistency and basic assumptions about Samatha meditation, which is not a strong constraint. In order to increase the credibility of the meditation model, transfer to other tasks would be necessary. Nevertheless, the positive transfer effect of the goal management operators to the SART indicates some valuable points. It suggests that the mechanisms of the meditation model are at least somewhat generalizable and are not merely artifacts of a specific modeling situation. It furthermore indicates that the mind-wandering paradigm, which was very similar in both models, is plausible. Furthermore, the transfer was congruent with the kind of change one would predict. What is more, the meditation model is quite robust, simple and produces reasonable behavior considering its parsimony. In other words, there is reason to believe that the model captures one important aspect that might underlie FA meditation: a feedback loop effect induced by patient and deliberate application of a goal management strategy. On the other hand, it does not capture aspects of meditation that reflect cultivation of a non-judgmental attitude and transformation of mental habits.

In short, we have presented the first computational model of meditation and have shown that it makes predictions for transfer to cognitive task performance. The model suggests that the transfer consists of goal management faculties and that it enhances performance through a feedback loop mechanism.

Acknowledgments

Mind & Life visiting scholarship to Marieke van Vugt. We thank Trudy Buwalda and Stefan Huijser for their help with *Prims*.

References

Anderson, J. R., & Lebiere, C. J. (2012). *The Atomic Components of Thought*. New York, NY, US: Psychology Press.

Desbordes, G., Gard, T., Hoge, E. A., Hölzel, B. K., Kerr, C., Lazar, S. W., . . . Vago, D. R. (2015). Moving beyond mindfulness: defining equanimity as an outcome measure in meditation and contemplative research. *Mindfulness, 6*(2), 356-372.

Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: Integrating mathematical and cognitive modeling. *Cognitive Science, 33*(5), 880-910.

Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy, 35*(4), 639-665.

Kabat-Zinn, J. (1990). *Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness*. New York, NY, US: Dell Publishing.

Khoury, B., Sharma, M., Rush, S. E., & Fournier, C. (2015). Mindfulness-based stress reduction for healthy individuals: a meta-analysis. *Journal of Psychosomatic Research, 78*(6), 519-528.

Lutz, A., Slagter, H. A., Dunne, J. D., & Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends in Cognitive Sciences, 12*(4), 163-169.

MacLean, K. A., Ferrer, E., Aichele, S. R., Bridwell, D. A., Zanesco, A. P., Jacobs, T. L., . . . Wallace, B. A. (2010). Intensive meditation training improves perceptual discrimination and sustained attention. *Psychological Science, 21*(6), 829-839.

Riley, D. (2004). Hatha yoga and the treatment of illness. *Alternative therapies in health and medicine, 10*(2), 20-21.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). "Oops!" performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35*(6), 747-758.

Sedlmeier, P., Eberth, J., Schwarz, M., Zimmermann, D., Haarig, F., Jaeger, S., & Kunze, S. (2012). The psychological effects of meditation: A meta-analysis. *Psychological Bulletin, 138*(6), 1139-1171.

Seli, P., Carriere, J. S. A., & Smilek, D. (2015). Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. *Psychological Research, 79*(5), 750-758.

Shear, J. (2006). *The Experience of Meditation: Experts Introduce the Major Systems*. St. Paul, MN, US: Paragon House.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, instruments, & computers, 31*(1), 137-149.

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120*(3), 439-471.

Teasdale, J. D., Segal, Z. V., Williams, J. M. G., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology, 68*(4), 615-623.

Vago, D. R., & Silbersweig, D. A. (2012). Self-awareness, self-regulation, and self-transcendence (S-ART): a framework for understanding the neurobiological mechanisms of mindfulness. *Frontiers in Human Neuroscience, 6*, 1-24.

van Vugt, M. K., Taatgen, N. A., Sackur, J., & Bastian, M. (2015). Modeling mind-wandering: a tool to better understand distraction. In N. A. Taatgen, M. K. van Vugt, J. P. Borst & K. Mehlhorn (Eds.), *Proceedings of the 13th International Conference on Cognitive Modeling* (pp. 1-6). Groningen, the Netherlands: University of Groningen.

Wallace, B. A. (1999). The Buddhist tradition of Samatha: Methods for refining and examining consciousness. *Journal of Consciousness Studies, 6*(2-3), 175-187.

Warm, J. S. (1980). The psychophysics of vigilance *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 24, pp. 605-605): Sage Publications.