NIMH RESEARCH DOMAIN CRITERIA (RDOC) PROJECT
WORKING MEMORY: WORKSHOP PROCEEDINGS

BETHESDA, MD
JULY 11-13, 2010

Background

The Research Domain Criteria (RDoC) project is designed to implement Strategy 1.4 of the NIMH Strategic Plan: "*Develop new ways of classifying disorders based on dimensions of observable behaviors and brain functions*." NIMH intends RDoC to serve as a research framework encouraging new approaches to research on mental disorders, in which fundamental dimensions that cut across traditional disorder categories are used as the basis for grouping patients in clinical studies. RDoC represents an inherently translational approach, considering psychopathology in terms of dysregulation and dysfunction in fundamental aspects of behavior as established through basic neuroscience and behavioral science research. The major RDoC framework consists of a matrix where the rows represent specified functional *Constructs*, summarizing data about a specified functional dimension of behavior, characterized in the aggregate by the genes, molecules, circuits, etc., responsible for it. Constructs are in turn grouped into higher-level *Domains* of functioning, reflecting contemporary knowledge about major systems of cognition, motivation, and social behavior. In its present form, there are five Domains in the RDoC matrix: Negative Valence Systems, Positive Valence Systems, Cognitive Systems, Systems for Social Processes, and Arousal/Regulatory Systems. The matrix columns specify *Units of Analysis* used to study the Constructs, and include genes, molecules, cells, circuits, physiology (e.g., heart rate or event-related potentials), behavior, and self-reports. The matrix also has a separate column to specify well-validated paradigms used in studying each Construct.

For detailed information about RDoC and the updated matrix, please see the RDoC web page.

To comment on any aspects of the proceedings, send email to the following: rdoc@mail.nih.gov.

The RDoC matrix is being developed to serve as a heuristic, and it is subject to change with scientific advances from the field. To "build the matrix," NIMH is bringing together leading experts to coalesce and articulate the state of knowledge for each of the five domains. Six meetings are planned in all; the first, reported below, details the proceedings of the workgroup addressing the working memory Construct within the Cognitive Systems Domain. This focus was chosen so the process could build upon the knowledge base organized by Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia Project (CNTRICS); further, it served as a pilot to test and develop the most efficient and productive format for the remaining five meetings to follow (one for each of the five major domains).

Workshop Proceedings

This workshop was the first in a series that will define Constructs, differentiate them from related Constructs, and examine what is known about the Units of Analysis for the Constructs within a Domain, what questions remain unanswered, and potential avenues of research that will answer these questions. We intentionally chose to discuss only one Construct, Working Memory (WM), selected from among those in the Cognitive Systems Domain; as noted previously, a future meeting will be devoted this entire Domain. The goals of this first workshop were to 1) arrive at an agreed-upon definition of WM, incorporating how the field views WM currently and how it is distinguished from other similar Constructs in cognition; and 2) provide an annotated listing (based on current knowledge) of the elements that would populate the RDoC matrix with respect to the genes, molecules, cells, and circuits sub-serving WM, as well as identifying promising behavioral tasks to reliably assess disruptions in WM functioning.

We invite and encourage comments on any aspect of the workshop and its outcomes summarized here; send comments via email to: rdoc@mail.nih.gov.

Note: This document is intended only as a summary of the workshop proceedings, and not as a definitive scientific review of the many issues involved. In this light, the references are meant only to be illustrative, and should not be considered as exhaustive or exclusionary.

Listed first is the definition that the workgroup developed, followed by a summary of the discussions leading to this definition.

Definition: Working Memory is the active maintenance and flexible updating of goal/task relevant information (items, goals, strategies, etc.) in a form that has limited capacity and resists interference. These representations: may involve flexible binding of representations; may be characterized by the absence of external support for the internally maintained representations; and are frequently temporary, though this may be due to ongoing interference.

- The participants agreed that WM is distinct from short-term memory (STM), long-term memory (LTM), cognitive control, attentional processes, and perception.
- Although participants agreed that WM and STM did represent different processes, there was no clear agreement regarding how they differed from each other.
- Differences between the concepts of WM and LTM centered around the fact that these processes appear to have distinct features: compared to WM, LTM has less accessible representations, a longer time frame, and may be reactivated/accessed more automatically; LTM is also associated with at least partially distinct circuits within the brain (medial temporal lobe versus dorsal frontal parietal in WM).
- Cognitive control processes appear to be distinguishable from WM by virtue of the fact that they involve processes that are not present in WM: motivation, goal selection, motor inhibition, and error conflict monitoring.
- WM differs from attentional processes in that attentional processes operate on or involve perceptual processes/representations and “bottom up inputs” that are not formally part of WM. In addition, the circuitry associated with the two concepts may be different: attentional circuits are generally considered more “parietal,” whereas WM circuits appear to be more “frontal.”
- Perception appears to be different from WM in that it depends on external versus internal support for the representations. It is possible that integrity of perceptual processes and representations may influence the integrity of derived WM representations, though more empirical work would be needed to establish this and perception does not appear to be a core component of WM per se.

Part One: Conceptual Aspects of Working Memory

The process followed by CNTRICS was a valuable guide for the NIMH RDoC team as they developed the draft matrix. Deanna Barch, PhD, and Cameron Carter, MD, who both lead the CNTRICS project, greatly helped the NIMH RDoC team in their role as moderators for this workshop. In their role as moderators, Drs. Barch and Carter began the first session by giving an overview of the open questions regarding theories underlying the concept of WM and how WM may differ from other cognitive processes.

Dr. Barch started the meeting with a discussion on three major questions that are relevant to theories regarding the concept of WM and its varying dimensions.

1. **Are elements of WM modality-specific?** Is it possible to have deficits in only some modalities of WM (e.g. visual/spatial versus language)?
2. **What is the nature of representations in WM?** Are there specialized WM representations or is WM a process of reactivating long term memory (LTM) representations?
3. **What is the concept of "capacity limitations" based on?**

1. **Are elements of WM modality-specific?**
   Are storage and maintenance of information carried out in dissociable modules, perhaps based on the nature of the information? One theory posits a central executive with visual/spatial WM buffers and verbal WM buffers; the latter are hypothesized to contain both phonological storage and an articulatory rehearsal loop. Data that support this theory include:
   - Single cell recordings demonstrate that different frontal neurons show activity when maintaining spatial versus object information;
   - Behavioral data demonstrate greater WM interference from a secondary task that uses the same modality or type of information that is being maintained in memory;
   - Neuropsychological data demonstrate that some patients appear to have selective impairments in visual/spatial versus verbal WM;
   - Neuroimaging data suggest that it is possible to activate somewhat different circuits based on the kind of information being maintained;
   - A large meta-analytical review of the WM literature (Wager & Smith, 2003) also supports this theory. The evidence reviewed suggests that WM processes are not limited to frontal areas of the cortex, but also include modality specificity in posterior regions, for example, parietal cortex.

   There are some aspects of this theory that are not resolved. For instance, when verbal rehearsal is subtracted out of imaging studies, much of the frontal activation (particularly inferior frontal) is removed, suggesting that the activation is not associated with storage during a WM task. It is also clear that a number of the mechanisms engaged in WM studies are also engaged by other cognitive processes. Finally, proposing separate buffers for every different types of information would create the need to postulate too many buffers (e.g., Postle, 2006). Furthermore, an alternative interpretation of the data presented above can support other theories: the meta-analytical review showed more evidence for modality specificity in posterior regions compared with frontal regions, and the data from single cell recordings can be interpreted as showing that neurons in frontal cortex are not necessarily modality-specific, depending on how the animal is trained or how aspects of the paradigm are modified (e.g., Cromer, Machon, & Miller, 2010). This brings us to the next question.

2. **What is the nature of representations in WM?**
   For reasons outlined above, a current alternative theory has been put forward that suggests that WM does not have modality-specific representations, but instead reactivates modality-specific representations from LTM. Findings and implications that support this theory include:
   - Studies that require manipulation of information in ways that do not involve WM (e.g., phonological processing or semantic processing) show activation in some of the same regions that are attributed to modality-specific WM buffers.
   - Meta-analyses and empirical studies have shown overlap in the neural circuits involved in WM and LTM. For example, maintaining images of houses or faces during a WM task is associated with

sustained activation in regions of the posterior cortex (where LTM representations are stored) that are specific to the modality of the information being maintained in the WM task.

However, this theory is still being debated, and in particular, three questions require more study:

- What mechanisms are responsible for encoding the temporal order of items in WM?
- Does the concept of rehearsal fully account for activation of representations in LTM?
- What other mechanisms may be involved with recruiting the specified regions associated with LTM recall? Is there a frontal buffer that binds WM with LTM representations?

3. **What is the basis for the concept of "capacity limitations" in WM?**
   Although it is clear from studies that capacity limitations reflect a major component of WM impairment in many forms of psychopathology, the neural bases of these limitations are still not fully understood. For instance, multi-store models suggest that limited capacity reflects the constraints on how rehearsal outpaces decay mechanisms (e.g. forgetting). In contrast, unitary store models suggest that limited capacity is a reflection of the constraints on attention.

   Three different potential mechanisms have been put forward in the literature to explain capacity limitations of WM.
   - Maintenance relies on synchronized oscillations in frontal or parietal regions that desynchronize over time due to variation in neuronal firing rates (passive decay). For successful maintenance of information, it is necessary that synchronized activity for different "items" be isolated from each other. Thus, capacity may be determined by the number of separate "activity bumps" that can be maintained separately at the same time (e.g., Lisman & Idiart, 1995; Comte et al., 2000; Jensen, 2006; Deiber et al., 2007).
   - The number of items that that can be maintained in working memory reflects the number of high frequency gamma cycles that can be embedded in a lower frequency data cycle (i.e. oscillatory activity for the different "items" may kept out of phase in order to be isolated from each other) (e.g., Lisman & Idiart, 1995; Lisman, 2010).
   - Capacity limits may be overcome, at least temporarily, by excitatory inputs into parietal cortex from prefrontal cortex. This capacity modulation could reflect the top-down control over WM capacity by the prefrontal cortex, and could explain individual differences or variation in the degree to which a person can use frontal mechanisms to modulate excitatory activity.

Dr. Carter followed with a review of the potential areas of overlap between WM and other cognitive systems, as well as conceptualizing these relationships. Cognitive systems are not orthogonal and thus it is reasonable to expect an overlap between components of WM and other cognitive systems. Three approaches to this overlap include:

1. <u>WM is a super-ordinate system that supports other cognitive systems and specific mechanisms in WM provide support for LTM.</u> For example, language comprehension is dependent upon WM capacity; this is reflected in individual differences in brain activity during language processing (e.g., Kaan & Swaab, 2002; Murray & Ranganath, 2007).
2. <u>WM is subordinate to other cognitive systems.</u> A recent meta-analysis, including data from over several hundred subjects completing subsets of classically defined executive functions, found that the frontal-parietal-cingulate network was consistently activated across all studies, suggesting that it may act as a general purpose cognitive control network supporting a whole range of different executive functions, including WM (Wager & Smith, 2003).
3. <u>WM and other cognitive functions form part of a distributive component process.</u> Studies suggest that the interrelationships between cognitive functions, including WM, are similar to different networks of

cerebral areas in both prefrontal and parietal regions. These networks are involved in motor planning, perceptual processing, and memory formation. They may be organized in a hierarchical way and, according to the task, different networks may work together to support task performance.

Part Two: Reaching Agreement on the Definition of Working Memory

The workshop participants met in two breakout groups. The charge for each group was to define WM, clarifying the boundaries as well as the interrelationships between WM, perception, attention, and LTM. Participants with expertise in genetics, cellular/molecular neuroscience, systems neuroscience, behavioral, and clinical sciences were represented in each of the breakout groups. Afterwards, the entire group reconvened to consider and compare the definitions from each breakout group. Elements common to the definitions proposed by each group were discussed and an agreed-upon definition of WM was proposed, as listed above.

Part Three: Populating the Units of Analysis in the Matrix for WM

The remainder of the workshop focused on listing specific elements, across the Units of Analysis, which have been reported in the literature as operational measures relevant to WM. The participants were also asked to identify critical issues needing further clarification for those elements, relevant to WM, which did not have sufficient evidence in the literature in order to complete the RDoC matrix with reference to WM. The breakout groups met to identify these elements, after which the full group reconvened for discussion. The discussion that followed centered on the necessity of finding mechanistically realistic links between the different units of analysis (genes, molecules, cells, neural circuits, behavior, and self-reports) that are relevant to WM.

The discussion about matrix elements resulted in an agreement that WM consists of four inter-related components, which are sufficiently distinct that each component requires its own set of entries for the Units of Analysis. These four components are active maintenance, flexible updating, limited capacity, and interference control. Accordingly, the matrix structure for this Construct is formatted to reflect these distinctions. Grant applications that involve WM may be directed at the overall Construct or at any of the four components whether singly or in combination.

Circuits

Participants were asked to develop a summary of how various circuits are thought to be involved in WM. Much research suggests that WM depends critically on sustained active maintenance in the prefrontal cortex (PFC), which interacts with posterior cortical areas in a task dependent manner. There are extensive loops from the PFC through the basal ganglia (striatum, substantia nigra pars reticulata (SNr), globus pallidus (GP), medial dorsal (MD) and ventral anterior (VA) thalamic nuclei) that may be important for driving the flexible updating of PFC representations, and in learning new tasks. There are connections between the PFC and medial temporal lobe that may support encoding of the contents of WM into LTM and retrieval of stored memories that can be reactivated in WM.

There may be gradients within these circuits along the following lines:

- A dorsal/ventral gradient related to representations of how versus what.
- A lateral/medial gradient related to representations of high versus low reward, value, etc.
- An anterior/posterior gradient related to representations of more versus less abstract information.

Activity in these areas is influenced by neuromodulators (e.g., dopamine and norepinephrine) that are released by neurons in midbrain areas (e.g., ventral tegmental area (VTA) and locus coeruleus (LC)) and play an important role in modulating the function of the PFC and basal ganglia. Dopamine and norepinephrine may influence the strength and/or fidelity of representations maintained in PFC. These systems influence many other aspects of functioning as well; as such, they are considered within the Arousal/Regulatory Domain in the RDoC matrix.

Behavior and Paradigms

The participants were charged with generating a list of tasks that were thought to best characterize the component processes of WM. They agreed that other WM tasks would likely be added to this list in the future. However, the group agreed that, to ensure construct validity in the measurement of WM, tasks would necessarily have to include the following essential task features.

1. In order to validate the task as an operational measure of WM and not another cognitive process (e.g., priming, episodic or habit learning systems) the tasks should:
    a. Require active maintenance of information:
        i. Information to be maintained should not be presented in the external environment; this will ensure that it is maintained "internally."
    b. The task should preclude (to the extent possible) allowing an individual to solve the task using a habit learning system.
        i. The task should require that the correct response to a given stimulus, OR that the stimuli themselves, vary across trials.
2. Many WM tasks do not involve specific manipulations of interference, although it is often assumed that interference is always occurring via the influence of previous stimulus traces, stimulus response mappings, competing task goals, or other information in the environment. However, a researcher could choose to explicitly manipulate interference control by the explicit presentation of competing information, goals, or task processes (e.g., either concurrently presented or from prior trials).
    a. This could take the form of having individuals perform some other type of processing while simultaneously holding other information in WM, such as in a range of complex span tasks designed by Engle and colleagues (Conway et al., 2005).

Potential Confounds:

- The workshop participants strongly encouraged investigators to ask: Is this task really measuring WM? What is the inference you wish to make about the source of the WM impairment?
- Simply showing deficits on a WM task may not be sufficient to confirm that the deficit actually reflects impairments in the core elements of WM. The purpose of these experimental tasks is to decompose their component processes in order to distinguish WM from cognitive processes such as perceptual encoding, various aspects of attention, fatigue, understanding task instructions, material specific processing deficits, visual/auditory acuity, and motivation.
- A number of tasks that have been used to measure WM meet one or more of the elements described above, and thus have reasonable construct validity as measures of WM. These have been included in the matrix below. However, it is important to note that this list is not meant to be exhaustive or exclusive.

NIMH encourages comments on any aspect of the workshop and the proceedings outlined here. Please send comments to: rdoc@mail.nih.gov.

Note: This document is intended only as a summary of the workshop proceedings, and not as a definitive scientific review of the many issues involved. In this light, the references below are meant only to be illustrative, and should not be considered as exhaustive or exclusionary.

REFERENCES

Comte, A., Brunel, N., Goldman-Rakic, P.S., & Wang, X-J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, *10*, 910-923.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769-786.

Cromer, J.A., Machon, M., & Miller, E.K. (2010). Representation of multiple, independent categories in the primate prefrontal cortex. *Neuron*, *66*, 796-807.

Deiber, M.P., Missonnier, P., Bertrand, O., Gold, G., Fazio-Costa, L., et al. (2007). Distinction between perceptual and attentional processing in working memory tasks: a study of phase-locked and induced oscillatory brain dynamics. *Journal of Cognitive Neuroscience*, *19*, 158-172.

Kaan, E., & Swaab, T.Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences*, *6*, 350-356.

Jensen, O. (2006). Maintenance of multiple working memory items by temporal segmentation. *Neuroscience*, *139*, 237-249.

Lisman, J.E. (2010). Working memory: the importance of theta and gamma oscillations. *Current Biology*, 20, R490-492.

Lisman, J.E., & Idiart, M.A.P. (1995). Storage of 7±2 short-term memories in oscillatory subcycles. *Science*, *267*, 1512-1515.

Murray, L.J., & Ranganath, 2007. The dorsolateral prefrontal cortex contributes to successful relational memory encoding. *Journal of Neuroscience*, *27*, 5515-5522.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*, 23-38.

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive, Affective, and Behavioral Neuroscience*, *3*(4), 255-274.

Participants

Woo-Kyoung Ahn, PhD, Yale University
Marie Banich, PhD, University of Colorado at Boulder
Helen Barbas, PhD, Boston University
Deanna Barch, PhD, Washington University
Robert Bilder, PhD, University of California, Los Angeles
Pamela Butler, PhD, New York University School of Medicine
Cameron Carter, MD, University of California, Davis
Ian Dobbins, PhD, Washington University
Randall Engle, PhD, Georgia Institute of Technology
Kimberly Andrews Espy, PhD, University of Nebraska-Lincoln
James Gold, PhD, University of Maryland School of Medicine
Raquel Gur, MD, PhD, University of Pennsylvania Medical Center

Ruben Gur, PhD, University of Pennsylvania Medical Center
Daniel C. Javitt, MD, PhD, Nathan Kline Institute for Psychiatric Research Professor of Psychiatry and Neuroscience, New York University School of Medicine
Angus Macdonald III, PhD, University of Minnesota
Earl Miller, PhD, Massachusetts Institute of Technology
Bita Moghaddam, PhD, University of Pittsburgh
Randall O'Reilly, PhD, University of Colorado at Boulder
Russell Poldrack, PhD, The University of Texas at Austin
J. Daniel Ragland, PhD, University of California, Davis, Imaging Research Center
Steven Silverstein, PhD, University of Medicine and Dentistry of New Jersey

NIMH RDoC Working Group Members

Bruce Cuthbert, PhD
Marjorie Garvey, MB, BCh
Robert Heinssen, PhD
Lucas Kempf, MD
Michael Kozak, PhD
Daniel Pine, MD
Kevin Quinn, PhD
Charles Sanislow, PhD, Wesleyan University
Phil Wang, MD, DrPH

NIMH Staff:

Thomas Insel, MD, Director, NIMH
Kathleen Anderson, PhD
Susan Borja, PhD
Emile Brouwers, PhD
Mi Hillefors, MD, PhD
Laurie Nadler, PhD
George Niederehe, PhD
Andrew Rossi, PhD
Steve Zalcman, MD
Varda Shoham, PhD
Alice Luo Clayton, PhD
Mary Ellen Oliveri, PhD
Scherri Jacobsen

Working Memory Matrix Specifications: Units of Analysis

"?" indicates some evidence, but not definitive

Unit of Analysis		Active Maintenance	Flexible Updating	Limited Capacity	Interference Control
Genes					
	NRG1/Neuregulin DISC1 DTNBP1/Dysbindin BDNF COMT DRD2 DAT1				
Molecules					
	Dopamine D1 D2	X X 	X X	X X (gain) ?	X X X
	Glutamate NMDA AMPA ?	X X	X	X	X
	GABA A ? B ?	X	?	X	X
Cells					
	Pyramidal	X			
	Distinct Types of Inhibitory Neurons Parvalbumin Calbindin Calretinin	X			X X X X
	Medium Spiny Neurons (Basal Ganglia)		X		
Circuits					
	Key Circuit: PFC-Parietal-Cingulate-Dorsal Thalamus-Dorsal Striatum				
	DLPFC		X	X	X
	VLPFC	X		X	
	Dorsal Striatum		X		
	Dorsal Parietal			X	
	Inferior Parietal	X		X	
	MD & VA Thalamus (by virtue of their role in circuit)		X	?	
Behavior and Paradigms					
	N-Back	X	X (?)	X	X (if you include non-target lures)

	Delayed Match to Sample	X	---	X	X (if you use repeated items, or delay period interference)
	Delayed Match to Non-Sample	X	----	X	X (if you use repeated items, or delay period interference)
	Sequence Encoding and Reproduction	X	----	X	----
	Sternberg Item Recognition (including recent negative variations)	X	X (recent negative task increases demand on updating)	X	X (if you use repeated items, recent negative variation)
	Complex Span Tasks	X	X	X	X
	Letter Memory/Running Memory	X	X (?)	X	X (?)
	Letter Number Sequencing	X	X	X	X
	Simple Span Tasks (may be more appropriate for developmental populations, in adults may not capture all key elements of WM)	X	------	X	X (if you use concurrent interference, as in Digit Span Distraction)
	Change Detection Tasks	X	----	X	------
	Keep Track Task	X	X	X	X
	AX-CPT/DPX	X	X	X	----
	Self-Ordered Pointing	X	X (?)	X	X
Physiology					
	Oscillations Gamma (local ensemble) Theta (local / thalamal cortical) Delta (cortico-cortical) ?				