

Supplementary Material S3: Coding Protocol

This document describes the coding procedures used to classify studies included in the systematic review *Working Memory Training Across the Senses*. The protocol ensured consistency in study inclusion/exclusion and data extraction.

The coding followed these six steps for each study selected for full-text screening and coding:

1. Assessment of study eligibility based on inclusion/exclusion criteria
2. Extraction of cognitive tasks used for training and pre-post assessment
3. Classification of cognitive tasks as WM or not WM
4. Categorization of WM tasks into family, paradigm and label
5. Classification of sensory modality of stimuli
6. Assessment of eligibility for cross-modal transfer

Coding procedures were defined prior to full data extraction and were applied consistently across all included studies. All coding and data extraction were performed by the first author using a structured spreadsheet. When uncertainties arose, cases were discussed among all authors and resolved through consensus.

The dataset is available in Supplementary Material S2 (full-text screening), S4 (study level dataset) and S5 (task level dataset).

1. Assessment of study eligibility based on inclusion/exclusion criteria

Study eligibility was assessed based on the full-text reports using the criteria in Table 1.

Table 1

Inclusion and exclusion criteria for studies in the systematic review.

Criterion	Inclusion	Exclusion
1. Language	English	Other
2. Population	Adult human participants	Non-human participants, children, or single-case/case series observations
3. Study type	Peer-reviewed empirical studies describing original experimental designs	Reviews, meta-analyses, theoretical papers, study protocols, or reports of experiments already reported in another included paper
4. Training design	Working memory training intervention with pre-post testing using working memory tasks	Interventions not targeting working memory, no pre-post testing, or assessing non-working memory outcomes only
5. Sensory modality	Description of sensory modality of stimuli used during training and testing	No information about sensory modality or no use of external sensory stimuli during training or testing

Any studies not meeting the criteria were marked as excluded with a comment about the reason for exclusion.

2. Extraction of cognitive tasks used for training and pre-post assessment

For each included study, all cognitive tasks used during training and pre-post training assessment were extracted from the methods sections of the articles.

For each task, the following information was recorded:

1. Name: A descriptive name to tell different tasks apart (e.g., visual n-back)
2. Study phase: Whether the task was used during training or assessment.
3. Implementation format: Whether the task was implemented as a standalone task or as part of a structured training battery (e.g., Cogmed), in which multiple tasks are integrated into a single intervention program. Standalone tasks could occur alongside other standalone tasks within the same study but were not part of an integrated training system.

If a study used multiple versions of the same task paradigm (e.g., visual and auditory n-back), each version was coded separately. Similarly, if the same task type was used for training and testing, they were coded separately.

3. Classification of cognitive tasks as WM or not WM

Each extracted cognitive task was classified as either a working memory (WM) task or a non-WM task.

A task was classified as a WM task if it required participants to maintain, manipulate, or update information over short periods of time in accordance with established WM paradigms described in the literature (e.g., Oberauer et al., 2018).

Tasks considered WM tasks included:

- n-back tasks
- span tasks involving manipulation or updating
- complex span tasks combining storage and processing
- tasks requiring maintenance of information for delayed recognition or recall

Tasks were classified as non-WM tasks if they primarily measured other cognitive processes such as:

- reasoning
- processing speed

- attentional control
- cognitive flexibility.

Forward span tasks were not classified as WM tasks because they are typically considered measures of short-term memory rather than working memory.

Although non-WM tasks were often included in the original studies as outcome measures, they were not included in the analysis of WM training modalities. If a study lacked WM tasks at either training or assessment, they were thus excluded from the review.

4. Categorization of WM tasks into family, paradigm and label

All tasks classified as single WM tasks were further categorized according to three hierarchical levels.

WM family: WM tasks were grouped into four functional families based on the primary cognitive process involved:

- maintenance – retention of information without transformation
- manipulation – transformation or reordering of stored information
- updating – continuous updating of stored information
- complex span – tasks combining storage with a concurrent processing task

WM paradigm: Within each family, tasks were classified according to the specific paradigm used. If a task did not fall within any paradigm, it was classified as “other”. This classification relied on the established WM paradigms described by Oberauer et al. (2018):

- maintenance: item recognition, change detection, relational recognition, probed recall, Brown-Peterson, continuous reproduction, reconstruction of order
- manipulation: memory updating (with transformation)
- updating: n-back, memory updating (with replacement), running memory span
- complex span

WM task label: To provide a more fine-grained characterization of tasks, an additional labeling level was introduced. The paradigms defined by Oberauer et al. (2018) capture broad functional categories but often group together task types that are treated as distinct in the literature. Therefore, when specific task types were sufficiently represented, they were assigned labels based on their conventional names (e.g., digit span, spatial span). These labels allowed differentiation between task variants within the same paradigm.

5. Classification of sensory modality of stimuli

Each WM task was coded according to the sensory modality of the stimuli presented during task performance.

The following modalities were included for coding: visual, auditory, olfactory, gustatory (not encountered), and tactile.

When stimuli from more than one modality were presented simultaneously, the task was classified as multimodal.

The classification was based on the modality of the stimuli presented to participants, not the modality of the response. Motor responses (e.g., button presses or reaching movements) were not classified as tactile stimuli. Likewise, auditory feedback sounds were not considered auditory task stimuli.

When modality was not explicitly reported, classification was determined using available methodological information such as stimulus descriptions, task instructions, figures, or referenced task protocols.

Studies were excluded if the modality of WM tasks could not be determined.

6. Assessment of eligibility for cross-modal transfer

Each study was evaluated to determine whether its design allowed assessment of cross-modal transfer.

Apart from the criteria in Table 1, a study was considered eligible for cross-modal transfer analysis if it met all of the following criteria:

1. Training was conducted using a single WM task per group.
2. Training was conducted using stimuli of a single sensory modality per group.
3. At pre- and post-training, performance was assessed on a training-equivalent task, meaning the same task used during training or an analogous task (e.g., adaptive and non-adaptive n-back).
4. At pre- and post-training, the training-equivalent assessment task was conducted within the sensory modality of the training as well as within an untrained sensory modality, separately.