Psychonomic Bulletin & Review
https://doi.org/10.3758/s13423-022-02144-7

THEORETICAL/REVIEW

Verbal interference paradigms: A systematic review investigating the role of language in cognition

Johanne S. K. Nedergaard[1] · Mikkel Wallentin[1] · Gary Lupyan[2]

Accepted: 29 June 2022

Abstract

This paper presents a systematic review of the empirical literature that uses dual-task interference methods for investigating the on-line involvement of language in various cognitive tasks. In these studies, participants perform some primary task X putatively recruiting linguistic resources while also engaging in a secondary, concurrent task. If performance on the primary task decreases under interference, there is evidence for language involvement in the primary task. We assessed studies (N = 101) reporting at least one experiment with verbal interference and at least one control task (either primary or secondary). We excluded papers with an explicitly clinical, neurological, or developmental focus. The primary tasks identified include categorization, memory, mental arithmetic, motor control, reasoning (verbal and visuospatial), task switching, theory of mind, visual change, and visuospatial integration and wayfinding. Overall, the present review found that internal language is likely to play a facilitative role in memory and categorization when items to be remembered or categorized have readily available labels, when inner speech can act as a form of behavioral self-cuing (inhibitory control, task set reminders, verbal strategy), and when inner speech is plausibly useful as "workspace," for example, for mental arithmetic. There is less evidence for the role of internal language in cross-modal integration, reasoning relying on a high degree of visual detail or items low on nameability, and theory of mind. We discuss potential pitfalls and suggestions for streamlining and improving the methodology.

Keywords Working memory · Dual-task performance · Language/memory interactions

Introduction

Does language help us think and solve problems, and if so, how? What kinds of mental tasks depend most on the use of language? These classic questions, debated in philosophy and psychology for more than a century (Fodor, 1975; Müller, 1978; Sokolov, 1968; Vygotsky, 1962; Watson, 1913), have been increasingly tackled using various empirical and modelling methods (Baldo et al., 2005; Coetzee et al., 2019; Feinmann, 2020; Gilbert et al., 2006; Luo et al., 2021; Romano et al., 2018). One widely used method is verbal interference or articulatory suppression (Perry & Lupyan, 2013). In studies using this method, participants are asked to perform some task that may or may not require linguistic processing while at the same time performing a clearly linguistic task, such as repeating a word. If performance on the "primary" task is compromised by the verbal task more than by control non-verbal tasks, one can conclude that language in some form is likely to be recruited by the primary task. Specific studies using this paradigm (e.g., Hermer-Vazquez et al., 1999; Newton & de Villiers, 2007) become held up as evidence for the crucial role of language as a cognitive tool (Bermúdez, 2003; Carruthers, 2002; Clark, 1998; Gomila et al., 2012). But follow-up studies and (non)replications complicate the narrative, and the use of different types of verbal interference and different types of control conditions makes comparisons across areas difficult. Finding that verbal interference disrupts one task but not another is difficult to interpret if the types of verbal interference that were used are substantially different.

Given the complexity, diversity, and potential importance of this literature, it is valuable to systematically review the findings to date. There exist reviews that focus on some domains where language has been proposed to play a role:

✉ Johanne S. K. Nedergaard
nedergaard@cc.au.dk

[1] Department of Linguistics, Cognitive Science and Semiotics, Aarhus University, Jens Chr. Skous Vej 2, 8000 Aarhus C, Denmark

[2] Department of Psychology, University of Wisconsin-Madison, Madison, WI, USA

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

Gilhooly (2005) for the role of language in reasoning when using verbal materials, Kiesel et al. (2010) and Koch et al. (2018) for the role of language in task switching, DeStefano and LeFevre (2004) and Raghubar et al. (2010) for the role of language in mental arithmetic, Ratliff and Newcombe (2008) for the role of language in spatial reorientation, and Alderson-Day and Fernyhough (2015) for a narrative review of the cognitive functions of *inner speech* specifically. Still lacking, however, is a comprehensive review across areas. This paper aims to provide a one-stop shop for dual-task evidence of the role of language in cognition. Importantly, dual-task approaches are just one way to investigate the role of language in cognition. Other ways include introducing new verbal labels as an experimental manipulation, examining performance by speakers of different languages, or attempting to interfere with linguistic processes with TMS (transcranial magnetic stimulation) or tDCS (transcranial direct current stimulation). Verbal interference remains a common method for testing on-line (i.e., in-the-moment) involvement of language in cognition, and so it is the method we focus on here.

Objectives

Our primary goals were:

1. To provide a coherent overview to aid in understanding of what cognitive functions language may and may not be involved in.
2. To provide suggestions and recommendations for methodology used in future studies in order to make results from different experiments more comparable.
3. To provide theoretically motivated reasons for choosing one interference type over another.

Verbal interference and verbal working memory

Verbal interference was first used in studying working memory (Baddeley & Hitch, 1974; Murray, 1967; Peterson, 1969), specifically to test the hypothesis that there is a component of working memory dedicated to the processing and storage of verbal material (the phonological loop and the phonological store) (Baddeley, 2003). Articulatory suppression (a type of minimally demanding verbal interference in which participants repeat a syllable or short word out loud) was used to discover whether participants were using verbal rehearsal to maintain the memory trace of for example a series of letters. The assumption that the phonological loop or verbal working memory is a specialized part of working memory underlies most of the studies reviewed here. We exclude studies specifically investigating this claim, but all the included studies rely on different verbal tasks drawing on the same resources, and thus that we have such cognitive components dedicated to processing in a verbal format – an assumption that has been called into question (Baddeley & Larsen, 2007; Jones et al., 2004, 2007). Criticism of the assumption revolves around whether verbal working memory is 'verbal' in an abstract sense or whether it simply involves low-level acoustic-articulatory processes. We omit discussion of this debate about the nature of "verbal" working memory because the logic of the dual-task design is valid regardless of the debate's outcome, even though it might be relevant when discussing how much of "language" different types of interference tasks plausibly interfere with.

In order to understand how verbal interference might work in more abstract cases, it is useful to first examine how it works in the most concrete, straightforward cases. Articulatory suppression has been used to investigate the so-called "phonological similarity effect" where serial recall performance is worse when the items to be remembered sound similar (Baddeley, 1966; Camos et al., 2013; Conrad, 1964; Conrad & Hull, 1964; Hintzman, 1967; Wickelgren, 1965a, b). The idea is that verbal working memory is divided into a phonological loop and a phonological store. Auditorily presented verbal material has direct access to the phonological store while verbal material presented visually (such as with written text) has to be converted in the phonological loop before it can enter the store. Thus, the phonological similarity effect should be different depending on presentation modality and the presence of articulatory suppression. See Fig. 1 for an illustration of an experiment testing the phonological similarity effect. Here, the hypothesis is that language is recruited to help store verbal material.

Because performing two tasks at the same time demands additional resources, performance under verbal interference must be compared to performance under an equivalently demanding but non-verbal dual-task condition. If verbal interference causes a more severe performance decrease than another distracting task equivalent in all other respects than the verbal, this would provide a causal argument for the presence of a linguistic component in the primary task. Articulatory suppression is often compared with the effect of foot tapping, another simple motor task that has been shown to be as attentionally demanding as articulatory suppression (Emerson & Miyake, 2003, Appendix A).

Outside the study of working memory components, verbal interference has been used to study, for example, task switching where the phonological loop is hypothesized to be recruited for self-cuing of whatever the relevant rule is, such as the common paradigm of switching between solving addition and subtraction problems. Here, verbal interference also impairs performance. In this specific case, the hypothesis would be that language is recruited to solve a task where it is necessary to maintain and update the relevant rule on each individual trial. This is similar to storing verbal materials in the phonological loop, except that

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

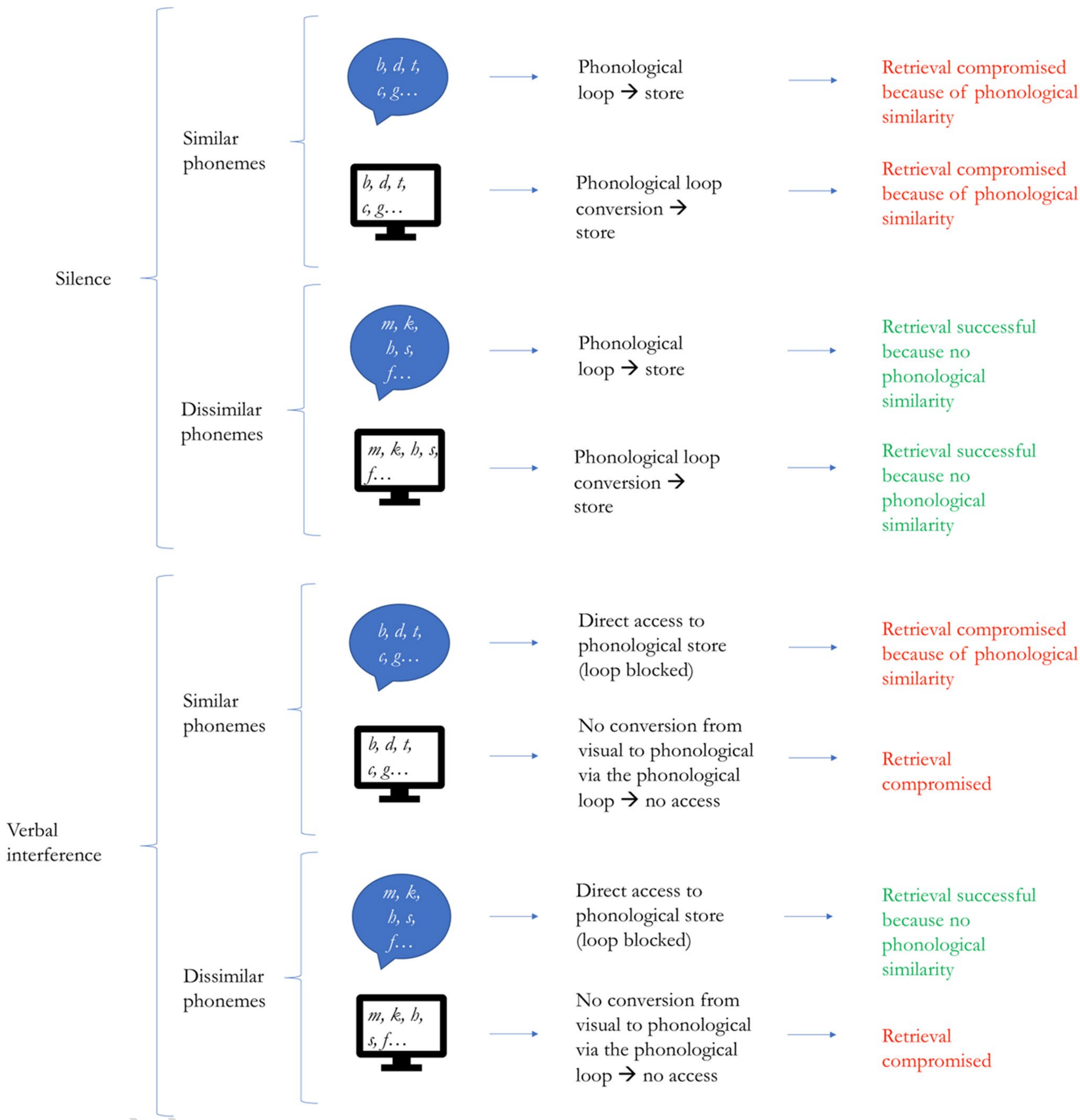

Fig. 1 A visualization of the mechanisms hypothesized to underlie the phonological similarity effect and how it differs depending on whether stimulus materials are presented verbally (speech bubble icon) or visually (screen icon)

instead of items to-be-remembered, the loop contains task instructions to-be-remembered. While covert language straightforwardly functions through verbal rehearsal in these examples, other studies have focused more on the structural and representational properties of language. These studies have used the dual-task interference methodology to test for example whether language aids cognition by providing the syntactic structure necessary for processing formal logic or by providing labelled categories that carve up otherwise continuous stimulus spaces. The precise mechanism of how repeating the word "December" (articulatory suppression) requires resources from the same cognitive component as recursive embedding and categorially labelled continua is less tangible than the precise mechanism of how articulatory suppression and task cuing might do the same. Similarly, many critics have pointed out the seeming paradox

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

of how language can have "deep" effects on non-verbal cognition that are nevertheless disrupted by surface-level verbal interference (Dessalegn & Landau, 2008; Gleitman & Papafragou, 2005; Li et al., 2009; see Lupyan, 2012a, b, for a discussion).

Verbal interference across cognitive domains

The more abstract, structure- or representation-focused dual-task studies are of a very different flavor compared with purely rehearsal-focused studies that have delineated the precise mechanisms and sub-mechanisms very precisely. There is, for example, a long way from testing whether the phonological similarity effect persists under articulatory suppression as illustrated in Fig. 1 (see e.g., Jones et al., 2004) to testing whether something like false belief understanding relies on covert language (see e.g., Newton & de Villiers, 2007). The hypothesis here could for example be that theory-of-mind processing requires on-line access to sentential complements (e.g., 'She thinks [the apple is in the box]') but how verbal interference would block this access is less clear as it has not been shown that participants have to formulate the sentence "she thinks the apple is in the box" explicitly in their minds to understand false belief on the fly. Thus, the easiest part of a study investigating the role of language in cognition with a dual-task experiment may be finding the effect – the more difficult part is explaining the precise mechanisms behind why this effect exists.

If there is one or several general roles that language plays in cognition, comparing the results of verbal interference across domains is one way of discovering what these might be. For example, most of the working memory-inspired studies included in the present review use very similar interference methods (word or syllable repetition) to test the role of covert language in task switching. By conducting slight variations on the primary task, these researchers thus zone in on whether covert language is recruited for task maintenance, task updating, task retrieval, etc. Once the precise effect is established, predictions are generated for other domains and we may test whether covert language also plays a role in for example task retrieval outside the addition/subtraction paradigm. Likewise, if we discover that verbal interference disrupts categorical perception of color, we should extend the paradigm to other types of categorical perception to ascertain whether covert language in general facilitates categorial perception. In the long term, it will of course also be necessary to integrate findings from other literatures apart from the dual-task interference literature (e.g., developmental evidence, evidence from brain lesions, evidence from noninvasive brain stimulation, etc.) As we proceed along this path, we can potentially map out domain-general functions of language for cognition, if such exist.

Review methodology

We followed PRISMA guidelines for selecting papers to include in this review (see Appendix B (OSM)). To be eligible, a paper needed to be peer-reviewed, and report at least one experiment with verbal interference and include at least one control task (either primary or secondary). Without such control tasks it is impossible to know whether the observed effects of verbal interference are purely due to the presence of a secondary task or whether they have something to do with language. We excluded studies in which the primary task being investigated was straightforwardly linguistic (e.g., lexical decision) because we were interested in the role of language in (putatively) non-verbal cognition. We also excluded papers with an explicitly clinical, neurological, or developmental focus. Although these studies are certainly valuable, including them would make it much more difficult to draw comparisons across areas, and so we leave their review for future work. We used the following search terms on PubMed and Google Scholar:

> *'articulatory suppression' OR 'dual-task paradigm' OR 'non-verbal control' OR 'verbal interference' NOT 'clinical' NOT 'developmental' NOT 'brain imaging'.*

To simplify the analysis of the findings, we divided the studies into clusters of primary task domains. If studies fitted into multiple clusters (e.g., if separate experiments within a study investigated different domains), the study is included in discussions of both clusters. For each study, the primary author collected the specific primary task(s), the specific interference task(s), the dependent variable(s), whether there was a selective effect of verbal interference, whether there was a difference between (levels of) the primary tasks, the number of participants in each experiment, and effect size(s) if reported. See Appendix A (OSM) for the full table including all the papers reviewed. The review was not registered, and a protocol was not prepared (aside from as detailed in the present section).

Results

Our literature search yielded 134 relevant papers, 33 of which were excluded (see criteria above), leaving 101 papers. We took great care to find as many of the relevant studies as possible, but as this literature is very fragmented and different subfields use different terminologies, we inevitably missed some. To the best of our knowledge, the present review represents an unbiased sample. We grouped the 101 relevant papers into 11 clusters based on the primary task: categorization (simple and complex), memory, mental arithmetic, motor control, reasoning (verbal and

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

non-verbal materials), task switching, theory of mind, visual change, and visuospatial integration and wayfinding. In the following sections, we discuss the findings of the systematic review in terms of both the types of interference tasks used and the cognitive functions investigated.

Types of interference tasks

The several different types of interference tasks present their own challenges. It is sometimes unclear whether an effect is simply due to irrelevant aspects of the interference tasks, and it is thus necessary to include them in our discussions and analyses. Aside from syllable or word repetition (n = 61), the main types of verbal interference used are verbal short-term memory tasks (n = 22), verbal shadowing (n = 13), and verbal judgment tasks (n = 6). Each of these types is discussed below.

Syllable/word repetition

Syllable or word repetition is by far the most common type of verbal interference used in the literature reviewed here, found in 61 of the 101 studies. This kind of verbal interference is often referred to as "articulatory suppression" because it suppresses normal function of articulatory organs. Syllable or word repetition were the only types of verbal interference found to be used to disrupt the role of covert language in task switching (Baddeley et al., 2001; Brown & Marsden, 1991; Liefooghe et al., 2005; Weywadt & Butler, 2013). For example, in Emerson and Miyake (2003) participants were asked to complete lists of alternating arithmetic problems while engaging in either repetition of the phrase "a-b-c" once every 750 ms or tap their foot once every 750 ms. The comparison interference task is either foot tapping, simple finger tapping, or pattern finger tapping. In experiments with more visually detailed primary tasks than the alternating lists paradigm, syllable repetition tends to be compared with both simple tapping and pattern tapping. Although there are also different ways of using this kind of articulatory suppression, the ways are plausibly comparable (i.e., there is no a priori reason to believe that repeating "the" twice per second would be different from repeating another short, well-learned word at the same rate). One study investigated whether the semantic content of the words being repeated mattered for a navigational working memory task (Piccardi et al., 2020). The experimenters asked participants to repeat nonsense syllables, egocentric spatial words, or non-egocentric spatial words, and this study found no difference between the different classes of words being repeated.

Verbal memory

Twenty-two studies reviewed here used a memory-based concurrent task (Annett & Leslie, 1996; Cheetham et al., 2012; Clearman et al., 2017; Croijmans et al., 2021; Frank et al., 2012; Gilbert et al., 2006, 2008; He et al., 2019; Hegarty et al., 2000; Imbo & LeFevre, 2010; Kranjec et al., 2014; Liu et al., 2008; Lupyan, 2009; Maddox et al., 2004; Newell et al., 2010; Robert & LeFevre, 2013; Samuel et al., 2019; Trbovich & LeFevre, 2003; Vogel et al., 2001; Winawer et al., 2007; Witzel & Gegenfurtner, 2011; Zeithamova & Maddox, 2007). In memory-based concurrent tasks, participants are asked to engage in covert rehearsal of verbal and non-verbal materials during the primary task with a subsequent memory test. For example, Lupyan (2009) investigated thematic or perceptual odd-one-out judgment with word or picture stimuli as the primary tasks and verbal and visuospatial memory as the secondary interference tasks. The interference tasks were either a nine-digit verbal rehearsal with a four-alternative forced choice test after each trial or a nine-dot spatial rehearsal with a four-alternative forced choice test after each trial. Another frequent version of this memory-based verbal interference task is N-back matching, where words are presented sequentially and participants have to press a button if a word matches the one immediately preceding it (Gilbert et al., 2006, 2008; Kranjec et al., 2014; Liu et al., 2008). One issue with using memory tasks as interference is that it is difficult to separate the different stages of memory encoding. If there are interference effects, it is difficult to see whether this happens at the encoding, maintenance, or retrieval stages. It could be that participants simply encode and store the to-be-remembered material outside working memory (e.g., in long-term memory) at the beginning of a trial, especially when trials last more than a few seconds. This enables them to devote all of their verbal resources to the primary task until they have to retrieve the to-be-remembered material again after the trial.

Verbal shadowing

In verbal shadowing, participants are asked to "shadow" continuous speech – i.e. repeat as quickly as possible without breaks – while simultaneously performing a primary task. Compared to syllable repetition, verbal shadowing has been used in a wider range of experiments. It was for example used in three of the four theory-of-mind experiments reviewed here (Dungan & Saxe, 2012; Forgeot d'Arc & Ramus, 2011; Newton & de Villiers, 2007), one of the memory studies (Perkins & McLaughlin Cook, 1990), one study on motion events (Feinmann, 2020), one study on categorization (Simons, 1996), one study on number representation (Frank et al., 2012), and six out of ten of the studies on visuospatial integration and wayfinding (Bek

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

et al., 2009, 2013; Hermer-Vazquez et al., 1999; Hupbach et al., 2007; Ratliff & Newcombe, 2005, 2008). For example, in Hermer-Vazquez et al. (1999), participants were asked to continuously shadow a tape recording of articles from a political newspaper. As a comparison interference task, Hermer-Vazquez et al. used a rhythm shadowing task where participants were asked to shadow-clap a sequence of clapped rhythm in 4/4 time that occurred at a rate of about 90 beats/min with a new rhythm played every eight beats. Rhythm shadowing is also used as the non-verbal interference task in the other studies using verbal shadowing.

The main difference between syllable repetition as discussed above and verbal shadowing is that verbal shadowing is arguably more demanding – to shadow successfully, you have to both perceive input and produce output at the same time. It is also less predictable and does not rely on overlearned sequences. Thus, the two verbal interference methods are not strictly comparable as verbal shadowing may target more aspects of natural language than simply the phonological loop.

Judgment tasks

Finally, six studies used judgment tasks as verbal interference, a more varied class of tasks that differ in their demands on response inhibition and comparisons between a presented stimulus and one (or several) held in memory. For example, Sims and Hegarty (1997) investigated "mental animation" (inferring the motions of mechanical systems) while having participants judge whether a specific letter was present in a list of six letters or not (putatively verbal interference) or decide if two patterns of four dots on a 4×4 grid were the same or different (a visuospatial interference condition). Hund (2016) and Meilinger et al. (2008) used similar interference tasks while examining wayfinding as the primary task. Here, the verbal interference task was word/non-word judgment. For the visual interference task, participants had to judge whether the two hands of the clock would be in the same half of the clock face or different halves of the clock face (dividing the clock face into an upper and a lower half) given a specific time of day (e.g., "6 o'clock"). Meilinger et al. (2008) also had a spatial interference task where participants were asked to judge from which direction a sound was coming. Pilling et al. (2003) used relative size discrimination and rhyme judgment. A special subclass of verbal judgment task is the Stroop task, where participants are presented with color words written with colored letters and have to respond based on the color of the letters and not the color name of the word. This type of judgment task was used in two studies, both testing motor control (Biese et al., 2019; Talarico et al., 2017).

Interim discussion of interference tasks

We found four main types of verbal interference tasks: syllable repetition, verbal memory, verbal shadowing, and judgment tasks. The review of the different tasks raised a few issues. First, it was not always clear to us which task was secondary and which was primary. Second, it is often difficult to assess performance on the interference task. Third, the verbal and non-verbal interference tasks do not always live up to the dual constraints of being (a) equally demanding and (b) different in only the presence or absence of "verbality" (Perry & Lupyan, 2013). We address these issues here.

In several studies we reviewed, it was unclear which was the 'primary' task and which was the 'secondary'. Usually, researchers are interested in investigating the role of covert language in a specific cognitive component which they term the primary task (e.g., memory for facial expressions) and use a secondary task (e.g., rhyme judgments) to interfere with the primary task. Many times, however, the distinction between primary and secondary task is merely a question of terms. Trying to memorize facial expressions might interfere with rhyme judgments, but making rhyme judgments might also interfere with trying to memorize facial expressions. It is necessary therefore to measure a potential trade-off effect where participants may devote all their resources to the secondary task instead of the primary task – if there is a trade-off effect, performance on the primary task and performance on the secondary task should be negatively correlated. Unfortunately, this is very rarely reported and often cannot be assessed because performance on the secondary task is generally not measured. This is, for example, the case with syllable repetition and verbal shadowing where the experimenters do not objectively assess performance, often simply writing something to the effect of: "The experimenters monitored that participants repeatedly uttered the word 'the' at 2 Hz." Without having some form of performance measure on the secondary task, we have no way of knowing how engaged participants are in the task, and whether the engagement fluctuates according to the demands of the primary task, for example, participants may strategically pause shadowing or verbal rehearsal when faced with a difficult trial on the primary task.

The third issue relates to how comparable the verbal and non-verbal interference tasks are. Ideally, the two tasks should be simultaneously equally difficult and attentionally demanding and differ only in their involvement of language. This is difficult to operationalize and has not always been done (or done well). Hermer-Vazquez et al. (1999), for example, ascertained that their verbal shadowing and rhythm shadowing tasks were equally demanding by assessing participants' performance on a visual search task and finding that the two interference tasks had comparable

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

detrimental effects. The conclusion that the two tasks are equally demanding in this case relies on the assumption that a visual search task would demand equal resources from verbal and visuospatial working memory, which is debatable. Relatedly with studies using syllable repetition, there has been some debate on whether the foot tapping task is an appropriate equivalent interference task in terms of demand. Proponents argue that it *is* equivalent because it is a simple motor task like repeating a word and should be as automatic and undemanding of the "central executive," the only difference between syllable repetition and foot tapping then being that syllable repetition involves articulatory organs (e.g., Emerson & Miyake, 2003, Appendix A).

In the discussions of the primary tasks investigated below, it is important to keep these interference task issues in mind. It may be the case that the presence or absence of verbal interference effects are not caused by the involvement or lack thereof of covert language but rather caused by incomparability of verbal and non-verbal interference tasks, hidden trade-off effects, or interference tasks that are not appropriate to the primary task investigated.

Effects of verbal interference on different cognitive tasks

We first describe the key studies from each family of primary functions we investigated and summarize the overall findings. The broad categories of primary functions investigated (ordered by how many studies each category contains) are: reasoning (verbal and non-verbal materials), memory, task switching, categorization (simple and complex), visuospatial integration and wayfinding, mental arithmetic, visual change, theory of mind, and motor control. See Appendix A (OSM) for a listing of the individual studies.

Reasoning

We identified 20 studies investigating reasoning. These can be divided into those using verbal materials (which encompasses studies that investigate formal logical problem-solving presented in a verbal format) and those using non-verbal materials (e.g., matrix reasoning, visual recursion, Tower of London).

Using verbal materials Eight studies investigated the role of covert language in reasoning using verbal materials (Evans & Brooks, 1981; Farmer et al., 1986; Gilhooly et al., 1993, 1999, 2002; Klauer, 1997; Meiser et al., 2001; Toms et al., 1993), which include propositional reasoning, conditional reasoning, and syllogistic reasoning. Here, covert language is hypothesized to help through providing a representational structure that facilitates reasoning with premises, conclusions, conditionals, assumptions, etc. Problems are presented in a verbal format and participants usually have to respond by saying whether the conclusion is valid or invalid.

Evans and Brooks (1981) tested participants on conditional reasoning and found that their rate of accepting invalid inferences was not affected by either simple, overlearned articulatory suppression (repeating the digits 1–6 in order) or articulatory suppression with a memory load (repeating the digits 1–6 in a random order specified by the experimenter). Somewhat surprisingly, response times were actually faster during articulatory suppression (this pattern is frequently seen; we comment on it in the Discussion). Testing both true/false judgments of declarative sentences about the order of two presented letters and mental rotation judgments, Farmer et al. (1986) found that digit repetition selectively impaired reasoning while spatial tapping selectively impaired the mental rotation judgments. In contrast with Evans and Brooks (1981), Toms et al. (1993) investigated conditional reasoning and found that articulatory suppression instantiated by repeating a simple overlearned sequence did not impair reasoning judgments, but that articulatory suppression with a memory load did. Specifically, the memory-load condition made participants less likely to accept valid modus tollens inferences (if p then q ➔ not q then not p). As Toms et al. (1993) themselves point out, there were some methodological differences between the two studies – most importantly, the study by Evans and Brooks used a between-subjects design, which could mean that it was not sufficiently sensitive to separate interference effects from individual differences in reasoning abilities.

Generally, the studies found a specific disruptive effect of random number generation but not of concurrent repetition of an overlearned sequence of digits. The latter was sometimes also disruptive – although the pattern is far from clear – but never more so than visuospatial concurrent tasks when these were included. Articulatory suppression seemed to be more disruptive when premises were presented sequentially than when they were presented simultaneously (see Gilhooly et al., 1993, 2002, respectively). Especially the finding that dual-task interference is observed with trained/skilled participants but not with untrained/low-skilled participants is relevant for the present review as an illustration of the idea that reliance on a verbal strategy in reasoning might depend on skill-level.

Using non-verbal materials Reasoning using non-verbal materials encompasses 12 studies, three of which included the Tower of London task as the primary task (Cheetham et al., 2012; Phillips et al., 1999; Wallace et al., 2017), two tested the Wisconsin Card Sorting Task (Baldo et al., 2005; Dunbar & Sussman, 1995), two tested a Visual Errands Test (Law et al., 2006, 2013), one tested paper folding, card rotations, and picture matching (Hegarty et al., 2000), one tested visual recursion (Martins et al., 2015), one tested the Hidden

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

Figures Test (Miyake et al., 2001), one tested Raven's Progressive Matrices (Rao & Baddeley, 2013), and one tested analogical mapping (Waltz et al., 2000). Generally, in these cases, language is hypothesized to be involved as a problem-solving tool where participants discuss with themselves or simulate potential solutions to the problems internally. It is also sometimes the case that covert language is hypothesized to help by providing a label for the rule when this has to be discovered (e.g., in the Wisconsin Card Sorting Task, in the Martins et al. visual recursion study, or in Raven's Progressive Matrices).

The Tower of London task requires participants to move a stack of discs from one peg to another while preserving a specific order (e.g., a smaller disc can never be under a larger disc). Of the three studies investigating the Tower of London task, only Wallace et al. (2017) found a specific effect of articulatory suppression with participants making more excess move in this condition. Cheetham et al. (2012) used memory-based interference tasks and found that only performance on the secondary tasks was affected – and not performance on the Tower of London task. Visuospatial memory was significantly worse when performed concurrently with the Tower of London task. Notably, Phillips et al. (1999) found that articulatory suppression had a *positive* effect on both completion time and error rate.

The Wisconsin Card Sorting Task requires participants to sort cards according to rules that they have to discover through trial-and-error and which change frequently. Dunbar and Sussman (1995) found a specific effect of articulatory suppression on perseverative errors (when participants persevere with sorting according to a rule that has changed) compared with tapping but no interference on number of categories achieved or non-perseverative errors. In contrast, Baldo et al. (2005) found that both articulatory suppression and foot tapping were associated with more perseverative and non-perseverative errors, but these two interference conditions were importantly not statistically different from each other. This means that we cannot say if the impairment was due to dual-task demands or specifically due to verbal demands.

The Visual Errands Test did not appear to be affected by verbal interference. In this kind of study, participants must complete a list of errands in a virtual environment while taking care not to break some rules. Thus, this task is more about planning and multitasking than about visuospatial orientation. In both studies (Law et al., 2006, 2013), the interference tasks hypothesized to involve the Central Executive (random month generation, tone localization) had larger negative impact than articulatory suppression. There was no specific effect of verbal interference on number of errands completed, number of errors/rule breaks, or time.

The remaining four studies in this section investigated the Hidden Figures Test (Miyake et al., 2001), Raven's Progressive Matrices (Rao & Baddeley, 2013), visual recursion (Martins et al., 2015), and analogical mapping (Waltz et al., 2000) respectively. The Hidden Figures Test is a visuospatial problem-solving test requiring participants to identify which of five simple figures is hidden inside a more complex figure. In Raven's Progressive Matrices, participants are presented with a set of patterns organized according to a specific rule, and need to figure out which of several patterns best completes a 3 × 3 matrix. In Martins et al. (2015)'s study, participants were asked to judge whether some visual patterns could be generated by recursive rules from other visual patterns. In the analogical mapping task investigated by Waltz et al. (2000), participants have to map visual scenes onto each other by their relational properties instead of their surface properties. None of these four studies showed a specific negative effect of verbal interference.

Taken together, verbal interference does not obviously disrupt visuospatial problem-solving of the kind tested in these studies. Only two of the 12 studies – Dunbar and Sussman (1995) and Wallace et al. (2017) – found a specific disruptive effect of verbal interference. Interestingly, in both Dunbar and Sussman (1995) and Wallace et al. (2017), verbal interference was associated with less inhibitory control, i.e., making more excess moves or continuing with perseverative errors. This may indicate that covert language is recruited for inhibitory control.

Memory

We found 17 studies that investigated memory under different interference conditions (Annett & Leslie, 1996; Brandimonte et al., 1992a, b; Croijmans et al., 2021; Gaillard et al., 2012; Gimenes et al., 2016; Henson et al., 2003; Hitch et al., 1995; Mitsuhashi et al., 2018; Nakabayashi & Burton, 2008; Pelizzon et al., 1999; Perkins & McLaughlin Cook, 1990; Souza & Skóra, 2017; Vandierendonck et al., 2004; Vogel et al., 2001; Walker & Cuthbert, 1998; Wickham & Swift, 2006). Covert language is hypothesized to aid memory in different ways, for example by providing a more abstract code for the item to be remembered in addition to the representation in the relevant sensory modality (Paivio, 1991). This is known as *dual coding theory* and posits that a memory trace is stronger if it is captured by both perceptual experience and verbal experience. Alternatively, covert language could aid memory by providing a medium for continuous rehearsal of the items to be remembered. Of course, these two hypotheses are not mutually exclusive as covert language could potentially aid memory both by encoding and by rehearsal.

Henson et al. (2003) did not find a specific detrimental effect of articulatory suppression on either a list probe task assessing memory for serial order of visually presented

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

letters or on an item probe task assessing memory for single item presence or absence. The three other interference tasks were irrelevant sound presentation, simple finger tapping, and complex, syncopated finger tapping. There was some indication that irrelevant sound and articulatory suppression had a larger detrimental effect on the list probe task than on the item probe task, although this was likely due to a ceiling effect on the item probe task. Thus, the results from Henson et al. (2003) do not support a selective role of covert language in either memory for either serial order or individual items. On the other hand, Nakabayashi and Burton (2008) reported a specific detrimental effect of articulatory suppression on facial recognition memory. Articulatory suppression during encoding was associated with worse performance on recognition memory compared with both a verbalization condition (where participants were asked to describe the faces out loud) and a simple tapping condition. Interestingly, Experiment 4 of Nakabayashi and Burton (2008) showed some indication that encoding the faces verbally *after* visual presentation had a weak detrimental effect on recognition memory. This suggests that the benefits of verbal encoding of visual stimuli depend on timing – this is reminiscent of the verbal overshadowing effect (Schooler & Engstler-Schooler, 1990), which is the finding that (forced) verbal descriptions of visual stimuli make subsequent recognition memory worse. In fact, Wickham and Swift (2006) investigated the verbal overshadowing effect specifically and found that verbal interference during stimulus presentation made the detrimental effect of subsequent verbal (over)description disappear.

Investigating memory for gestures, Gimenes et al. (2016) found that a verbal strategy (training manipulation) for remembering gestures was better than a gestural strategy, and that verbal interference interfered with gesture reproduction accuracy regardless of strategy. In a similar study, Mitsuhashi et al. (2018) found a specific effect of verbal interference on the Luria Hand Test, which measures reproduction accuracy. Less conclusive evidence for the facilitative role of language in memory comes from Walker and Cuthbert (1998), who investigated memory for color-shape associations, only using articulatory suppression as an interference task – thus it is not possible in this case to tell if there was a specific effect or not. However, they found that articulatory suppression disrupted the nameability advantage associated with some of the stimuli, supporting the idea that linguistic labelling facilitates memory. Interestingly, Souza and Skóra (2017) also found that overtly labelling colors to be remembered facilitated reproduction accuracy but also made the memory representation more categorical – in contrast, concurrent syllable repetition had a detrimental effect on reproduction accuracy.

Four of the memory studies tested the effect of verbal interference on both recognition memory and mental transformations of images (Brandimonte et al., 1992a, b; Hitch et al., 1995; Pelizzon et al., 1999). These studies found that while verbal interference disrupted recognition memory, mental transformation of the images to be remembered was actually improved by verbal interference. Mental transformation in this case refers to subtracting elements from the images, rotating them, or combining them to produce other recognizable forms. In addition, both advantages and disadvantages (e.g., stemming from degree of nameability) associated with verbal labelling disappeared with verbal interference. The authors of these four studies interpret the findings to mean that we normally use verbal resources to name visual stimuli to be remembered, and that this helps us recognize the stimuli later. However, the stored representation in verbal format does not maintain all the details of the original visual stimuli, which is why manipulations that depend on visual details are easier under verbal interference. This interpretation fits well with the color memory study by Souza and Skóra (2017) discussed above.

In most memory studies, the material to be remembered is presented visually, and nameability effects are found. However, some studies have also investigated the olfactory modality and memory for odors. Olfactory memory has been argued to depend on both a verbal code (taking advantage of odor labels) and a visual code (encoding an odor as the image of an object that prototypically smells like that). In a study that tested memory for wine odors, Croijmans et al. (2021) found that while experts were better than novices at both recognition and free recall, verbal interference had no effect on either group. Of the two other olfactory memory studies, one also did not find that verbal interference negatively affected memory performance (Annett & Leslie, 1996) and one found that digit shadowing had a specific negative effect on recognition, but not free recall (Perkins & McLaughlin Cook, 1990). Thus, there is no firm support for the on-line role of covert language in olfactory memory.

In summary, encoding items to be remembered verbally can be both beneficial (e.g., nameability advantages) and detrimental (e.g., verbal overshadowing effect), depending on what is to be remembered. The studies discussed here appear to support the idea that covert language influences memory as both advantageous and disadvantageous effects associated with verbal encoding disappeared under verbal interference.

Task switching

The present review found 16 studies investigating the role of covert language in task switching (Baddeley et al., 2001; Brown & Marsden, 1991; Bryck & Mayr, 2005; Emerson & Miyake, 2003; Grange, 2013; Kirkham et al., 2012; Liefooghe et al., 2005; Miyake et al., 2004; Saeki, 2007; Saeki et al., 2006, 2013; Saeki & Saito, 2004a, b, 2009;

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

Tullett & Inzlicht, 2010; Weywadt & Butler, 2013). All these studies test participants' ability to switch between two tasks and measure switch cost on reaction time and error rate (i.e., how much slower are the responses when a task B trial immediately follows a task A trial compared to if it follows another task B trial). These tasks included adding and subtracting numbers (Baddeley et al., 2001; Emerson & Miyake, 2003; Saeki & Saito, 2004a), color or shape sorting tasks (Kirkham et al., 2012; Liefooghe et al., 2005; Miyake et al., 2004), numerical or physical size judgment tasks (Saeki, 2007; Saeki et al., 2006, 2013; Saeki & Saito, 2004b, 2009), a Stroop task (Brown & Marsden, 1991), arithmetic problems verification (Bryck & Mayr, 2005), detection of different visual shapes preceded by visual cues (Grange, 2013), switched and regular versions of a Go/No-go task (Tullett & Inzlicht, 2010), and voluntary switching between odd/even and high/low digit judgments (Weywadt & Butler, 2013). It is worth noting that it is difficult to say if these task-switching experiments investigate flexibility (as participants need to flexibly shift between task sets) or inhibition (as participants need to inhibit the responses that they would make according to the non-active task set), or indeed if these two processes are two sides of the same coin.

As is evident from the above list, there are several different types of switch tasks represented in this primary task category – however, they all have in common that participants are asked to switch between responding to the same stimuli according to the rules of two different task sets. Usually, the studies also compare conditions where the relevant rule is somehow cued (e.g., displaying a '+' when the task is to add and a '−' when the task is to subtract) to conditions where the relevant rule is not cued or cued in a different way (e.g., endogenously vs. exogenously). Participants are hypothesized to retrieve and maintain the relevant rule or task set verbally. When the relevant rule is externally cued, articulatory suppression should have no effect if verbal rehearsal is under normal circumstances used as a sort of internal cue. Additionally, the studies also all use syllable repetition and foot or finger tapping as verbal and non-verbal interference tasks.

As an example of one of these task-switching studies, Baddeley et al. (2001) conducted seven experiments where they varied the types of interference task while participants completed either blocked or switched lists of numbers to be added or subtracted. The task on an individual trial either required the participant to remember the rule (endogenous condition) or included the rule as indicated by a plus or a minus sign (exogenous condition). Performance on switched trial lists was slower than on blocked trial lists – the experimenters measured the cumulative reaction time on a list where the participants had to alternate between adding and subtracting 1 and a list where they always had to either add or subtract 1. There were two different interference tasks as well: articulatory suppression (reciting days of the week or months of the year) and task taxing the central executive *and* verbal working memory (alternating day of the week and month of the year; Monday – January – Tuesday – February etc.). The executive task was associated with slower performance on both switched and blocked trials while articulatory suppression only appeared to slow performance on switched trials. Further, reaction times were slower with verbal interference on endogenously versus on exogenously cued trials. This difference between reaction times presumably indicates the cost associated with maintaining and drawing on a mental representation of the task (adding or subtracting).

Overall, the pattern of results from these 16 studies supports the idea that covert language is used to retrieve and maintain the task-relevant rule. Articulatory suppression seems to disrupt task switching when task cues are not present in the stimuli (Emerson & Miyake, 2003), suggesting that verbal rehearsal is needed to "remind" the participant of the task at hand.

Categorization

Sixteen studies investigated the role of language in categorization (Gilbert et al., 2006, 2008; He et al., 2019; Liu et al., 2008; Lupyan, 2009; Maddox et al., 2004; Minda et al., 2008; Newell et al., 2010; Pilling et al., 2003; Roberson & Davidoff, 2000; Winawer et al., 2007; Witzel & Gegenfurtner, 2011; Zeithamova & Maddox, 2007). In categorization studies, covert language is hypothesized to aid cognition by providing labels to carve up continuous perceptual space, for example, the color spectrum (Lupyan, 2012a). In studies that investigate novel category learning, covert language is supposedly recruited for learning discrimination patterns that are rule-based and easily verbalizable. In contrast, discrimination patterns that rely on more high-dimensional patterns are hypothesized to be learned in a more procedural way (see e.g., Maddox & Ashby, 2004). There are important differences between studies where participants need to categorize along some criterion (e.g., that does not belong based on size) and odd-one-out/perceptual matching studies. These tasks vary a great deal in how much you need to know to perform well, for example, detecting a visual difference versus using semantic knowledge or learned rules to solve a given categorization problem. Therefore, we divide this section into "simple categorization" and "complex categorization." The first section includes studies investigating perceptual discrimination and matching within and between known categories. The second section includes studies that involve learning novel categories and forming ad hoc categories involving, for example, focusing on one dimension while abstracting over other dimensions.

Simple categorization These studies investigate the use of already existing categories for detection of differences (e.g., between different colors). Most of them focus on color categories, although the categorization of facial expressions, spatial relations, and animals have also been investigated. In the color classification studies, participants are presented with a color and asked to classify it or presented with a selection of colors and asked to find the odd one out. In Gilbert et al. (2006), for example, participants were presented with a circle of colored squares where all except one were the same color. Participants then had to respond indicating which half of the circle the odd colored square was in. The color of the odd square was either in the same color category as the remaining squares (e.g., a different shade of green) or in a different color category (e.g., blue among greens). This study found that there was a cross-category advantage in the right visual field, possibly related to verbal labels, but that this advantage disappeared under verbal interference. A later study, however, attempted to replicate the Gilbert et al. (2006) findings but found that if the colors were more carefully controlled, the effect of visual field disappeared and did not differ depending on the presence or absence of verbal interference (Witzel & Gegenfurtner, 2011). Other studies without verbal interference have successfully replicated the visual field effect (Zhong et al., 2015; Zhou et al., 2010). In a study testing Russian- and English-speaking participants, Winawer et al. (2007) found the two groups differed when they were asked to discriminate shades of blue that were either within-category or across-category for the Russian speakers (Russian "blue" is divided into two separate terms, "goluboy" meaning lighter blues and "siniy" meaning darker blues). There was a category advantage for Russian speakers but not for English speakers. The Russian category advantage disappeared with verbal interference. A parallel effect was found by He et al. (2019), who tested Chinese and Mongolian speakers (the latter have different color words for light blue and dark blue, the former do not). Extending the category effects found in color discrimination, Gilbert et al. (2008) investigated categorization of dog and cat silhouettes and found that the language-based categorization effect was stronger in the right visual field than in the left, and that this category effect was attenuated by verbal interference.

Kranjec et al. (2014) tested categorical and coordinate spatial relation tasks and found that a one-back word-matching task had a larger disruptive effect than a one-back pattern-matching task. In these spatial relations tasks, participants were asked to make same/different judgments of dot-cross configurations that differed in how verbalizable the differences were. Counter to the author's prediction, there was no difference between the effect of verbal interference on trials with easier-to-name versus harder-to-name spatial categories. Two other studies investigating categorical and coordinate spatial relation tasks did not find specific effects of verbal interference (Dent, 2009; van der Ham & Borst, 2011). These two both used syllable repetition as the interference task, although only one (van der Ham & Borst, 2011) also included a non-verbal interference task (finger tapping).

Investigating categorical perception of both color and faces, Roberson and Davidoff (2000) found a selective interference effect of a verbal concurrent task. With the verbal concurrent task, the increased accuracy usually associated with cross-category judgments relative to within-category judgments had disappeared. The authors interpret this as indicating that the advantages associated with categorical perception and memory of faces and colors derive from verbal encoding and storage. In an attempt to replicate Roberson and Davidoff's (2000) experiment, Pilling et al. (2003) found that if the type of interference task was unpredictable, the category advantage survived verbal interference. The authors suggest that unpredictability of interference task condition may have discouraged the use of a verbal strategy. In another study that similarly calls into question the role of on-line language in categorical perception of color, Liu et al. (2008) found that the cross-category boundary advantage survived verbal interference. Although these studies show somewhat conflicting results, they indicate some tentative support overall for the idea that linguistic labels facilitate the speed and accuracy with which we make discrimination and detection judgments.

Complex categorization In one group of studies, participants are asked to learn novel categories where the category structure is either rule-based and easily verbalizable (e.g., "red things are in category A, blue things are in category B") or where the category structure relies on information-integration (where at least two differently expressed dimensions need to be combined) and is not easily verbalizable. Support for this distinction comes for example from Maddox et al. (2004), who found that a four-digit memory task disrupted the learning of rule-based category structures but not information-integration category structures. Similarly, Minda, Desroches, and Church (2008) found that adults under verbal interference displayed a category-learning pattern similar to that of children in that they found disjunctive rules harder to learn ("red and small OR blue and large things are in category A, blue and small things OR red and large things are in category B"). Zeithamova and Maddox (2007) found that both a visual and a verbal concurrent memory task disrupted rule-based category learning but not information-integration category learning. In interpreting the results of these studies, it is important to take into account that Newell et al. (2010) found that the dissociation between information-integration and rule-based categorization disappeared when only participants who actually learned the rule were included in the analysis.

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

In a study investigating complex processing of already learned category structures, Lupyan (2009) investigated effects of verbal and visuospatial interference on participants' ability to appreciate different kinds of similarities among pictures of familiar objects (or words denoting those objects). Participants were shown three pictures or words and asked to choose the object/word that was most different from the two based on its real-world color, size, or thematic/function relationship. The study was based on prior work showing that individuals with aphasia were selectively impaired when asked to isolate specific perceptual dimensions such as color or size, but were similar to controls when asked to group on more thematic or functional criteria (Cohen et al., 1980; Davidoff & Roberson, 2004; De Renzi & Spinnler, 1967; see Vignolo, 1999, for review). Lupyan sought to determine whether a similar dissociation could be observed in non-aphasia participants whose language was interfered with during the task, and found that verbal interference selectively affected color and size trials for both picture and word stimuli.

Visuospatial integration and wayfinding

Twelve studies investigated the role of covert language in visuospatial integration and wayfinding (Bek et al., 2009, 2013; Caffò et al., 2011; Garden et al., 2002; Hermer-Vazquez et al., 1999; Hund, 2016; Hupbach et al., 2007; Labate et al., 2014; Meilinger et al., 2008; Piccardi et al., 2020; Ratliff & Newcombe, 2005, 2008). In these studies, covert language is supposed to help by providing a common medium for the integration of information from different sensory modalities as well as different types of information from the same sensory modality (e.g., shape and color).

Hermer-Vazquez et al. (1999) is one of the most famous studies in this field and widely cited in philosophy of cognitive science as evidence for the role of language in cognition (Carruthers, 2002; Clark, 1998; Gomila et al., 2012). In the original study, participants were placed in a rectangular room and saw something being hidden in one of the corners of the room. They were then blindfolded and spun around until they were thoroughly disoriented. The dependent variable in this kind of study is participants' search behavior – which corner do they search in? How do they reorient themselves? Originally, Hermer-Vazquez et al. (1999) found that participants engaged in verbal shadowing were unable to combine geometric and color features of the room to find the right corner (i.e., using both the fact that two walls were shorter than the others and the fact that one end wall was painted a different color).

Six of the remaining studies reviewed include attempts to replicate and extend these findings, unsuccessfully in all cases. To test whether the size of the room mattered, both Hupbach et al. (2007) and Ratliff and Newcombe (2008; Experiment 3) used a bigger room than Hermer-Vazquez et al. (1999), and found that only a spatial interference task impaired reorientation performance. Bek et al. (2009) compared prose shadowing and syllable shadowing and found that neither reduced performance to chance levels as in Hermer-Vazquez et al. (1999). Testing the effect of the specific instructions given to participants, Ratliff and Newcombe (2005) tested the difference between implicit and explicit directions and found no specific effect of verbal interference. Similarly, Bek et al. (2013) found that prose and syllable shadowing both only disrupted reorientation performance when instructions were vague and non-specific like in Hermer-Vazquez et al. (1999). There was no difference between the two shadowing types. Further variations of the original paradigm include a study by Caffò et al. (2011) that tested a virtual version of the reorientation task with syllable repetition as the verbal interference task and spatial tapping as the spatial interference task. Performance during both interference tasks was worse than the control condition, but spatial interference was significantly worse than verbal interference. There is a risk, however, that this was a motor artifact – participants had to perform spatial tapping with the left hand and navigate the virtual environment with a joystick with the right hand.

The remaining five experiments in this category investigated wayfinding in various more complex ways. Labate et al. (2014) examined learning of maps including landmarks and routes through navigation in a real environment and found that a spatial tapping task was worse for performance than a syllable repetition task. Comparable results were found by Meilinger et al. (2008) and Hund (2016), who investigated similar wayfinding tasks with similar interference tasks, namely word/non-word judgments as the verbal interference and clock hand judgments as the visual interference. Both studies found that the visuospatial interference tasks had a stronger detrimental effect on performance than the verbal interference tasks. Potentially shedding light on the different contributions of visuospatial and verbal working memory, Garden et al. (2002: Experiment 2) found that AQ3 the degree to which participants were affected by verbal and visuospatial interference tasks in a real-world navigation problem depended on individual differences in spatial ability. Specifically, participants with high spatial ability were more affected by a concurrent spatial tapping task, and conversely participants with low spatial ability were more affected by a concurrent verbal interference task. Further testing the effect of many different kinds of interference tasks, Piccardi et al. (2020) investigated navigational working memory and found that only sound localization disrupted performance. The other interference tasks were stationary walking, stationary complex movements, nonsense syllable

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

repetition, repetition of egocentric spatial words, and repetition of non-egocentric spatial words.

Despite early findings, the studies discussed in this section taken together do not provide strong support for the idea that covert language is recruited for visuospatial integration and wayfinding.

Mental arithmetic

Nine studies investigated cognitive processes related to mental arithmetic and exact number representation (Clearman et al., 2017; Frank et al., 2012; Imbo & LeFevre, 2010; Lee & Kang, 2002; Logie et al., 1994; Robert & LeFevre, 2013; Seitz & Schumann-Hengsteler, 2000, 2002; Trbovich & LeFevre, 2003). The phonological loop is hypothesized to help with mental arithmetic by keeping track of partial results needed for further computations (Ashcraft, 1995; Imbo et al., 2005). The studies often contrast arithmetic problems that require fact retrieval (usually small problems < 10) and problems that require carry operations. Most of the studies in this section found that verbal interference disrupts mental arithmetic across varying presentation formats (auditorily, visually, horizontally, vertically), problem size, and kind of mental arithmetic (addition, subtraction, multiplication). However, testing the effect of different distractors, Clearman et al. (2017) found that attending to the color and location of three dots for subsequent recall had a larger adverse effect on the speed of mental arithmetic than attending to words presented aurally for subsequent recall. Thus, there was no evidence of specific verbal involvement. Frank et al. (2012), on the other hand, found that both verbal shadowing and a memory task disrupted exact number representation for larger quantities. They conducted three experiments, only one of which included a control interference task – a comparison between memory for a sequence of consonants and a sequence of dot locations on a grid. Taken together, these studies seem to indicate that covert language resources are recruited for mental arithmetic problems that are most effectively solved using a verbal code – this includes problems featuring carry and borrow operations, problems presented horizontally (contrasting with vertically presented problems that appear to invite visual strategies), and problems presented auditorily.

Visual change

The six studies in this category include those investigating visual change detection (Hollingworth, 2003; Sense et al., 2017; Simons, 1996), mental animation (Sims & Hegarty, 1997), similarity ratings of motion events (Feinmann, 2020), and visuospatial construction and memory (Bek et al., 2009: Experiment 1). Bek et al. (2009) found a specific detrimental effect of verbal interference, but this effect was limited to one of their tasks. They used a block design task in which participants were asked to construct two-dimensional designs of red and white blocks, and a complex figure task in which participants were asked to copy a figure and draw it again from memory after a delay. Verbal shadowing only interfered with the complex figure task and only if participants were shadowing during the encoding stage and not the retrieval stage. The authors argue that the reason verbal shadowing interfered with the complex figure task and not the block design was that the complex figure task contained nameable elements. Nameability was also an important factor in Simons (1996) where the advantage associated with change detection for common objects (hats, chairs, etc.) disappeared with verbal shadowing. Interestingly, Hollingworth (2003) compared detection of rotation change and token change and found that token change detection was in fact more accurate with verbal interference than in a control condition.

Theory of mind

Four studies have investigated the on-line role of covert language in theory of mind (Dungan & Saxe, 2012; Forgeot d'Arc & Ramus, 2011; Newton & de Villiers, 2007; Samuel et al., 2019). Theory of mind refers to the ability to attribute thoughts, beliefs, intentions, etc. to other humans, even when these are at odds with one's own worldview. The connection between language and theory of mind is a much debated topic with input from developmental psychology (Lohmann & Tomasello, 2003), evolutionary psychology (Dunbar, 1998; Malle, 2002), and neuroscience (Siegal & Varley, 2006), among others. One hypothesis for why language would aid theory of mind is that the syntactic structure of sentential complements is recruited for representing other people's mental states, for example, "she thinks [that the apple is in the box]" (de Villiers, 2007; de Villiers & de Villiers, 2000; de Villiers & Pyers, 2002). Alternatively, the connection between theory of mind and language in development could be that hearing adults talk about mental states directs children's attention to unseen mental states as well as the abstract properties that superficially different situations have in common (Milligan et al., 2007).

Of the four studies reviewed here, only Newton and de Villiers (2007) found a specific effect of verbal interference on a theory-of-mind task where participants were asked to choose the correct ending for false belief videos. There was no effect of either verbal shadowing or rhythm shadowing (the comparison task) on true-belief videos. There are some issues with this experiment, however. For example, the authors did not include a control condition with no interference or attempt to equate the two interference tasks for difficulty. This latter point was rectified by Dungan and Saxe (2012), who found that when the verbal and non-verbal interference conditions were better

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

equated for difficulty, there was no effect of verbal interference on false belief reasoning. Similarly, Forgeot d'Arc and Ramus (2011) compared belief judgment tasks and mechanistic judgment tasks, and found that verbal shadowing had an overall effect on performance but not specifically on belief attribution. They did not compare with another interference task. Testing the effect of a different type of verbal interference task, Samuel et al. (2019) compared performance on false belief and false-photograph trials with interference tasks that involved an eight-digit covert rehearsal with a memory test and a 4 × 4 grid pattern rehearsal with a memory test. This study did not find that the false belief task was specifically impaired by the verbal interference task. It is worth noting that the interference here was not during the encoding stage but instead between encoding and retrieval. Nevertheless, the results of these four studies seem to indicate that there is little evidence that covert language is involved in on-line theory-of-mind reasoning.

Motor control

We found two studies that investigated the role of covert language in motor control in some way: jump landing performance (Biese et al., 2019) and single leg postural control (Talarico et al., 2017). The reasoning behind why covert language would help with motor control stems from Vygotskian self-regulation, according to which we use our inner voice to control our own behavior (Vygotsky, 1962). Covert language focuses attention on motor control and can be used to cue specific subcomponent motor actions that facilitate the overall movement goal (e.g., jumping, serving, hitting, etc.). Both studies found that a verbal interference task had a specific disruptive effect, one on reaction time (Biese et al., 2019) and one on squatting speed and depth (Talarico et al., 2017). Both studies compared physical performance during a Stroop Color Word test versus on a Brooks Visuospatial task, but these two interference tasks are not necessarily equated in other respects than the verbal (see Judgment tasks section above). This lack of comparability is underscored by the fact that both the Stroop Color Word test and a Symbol Digit Modalities test (basically an association memory test) had adverse effects on jump landing performance in Biese et al.'s (2019) study. Thus, there is some doubt as to whether it was the verbal component of the Stroop task that caused the interference or just attentional demands – the Stroop task also is not "pure" verbal interference in that sense as it also puts demands on executive control (response inhibition).

Discussion

As the above review has illustrated, the literature investigating the role of covert language in cognition using dual-task methodologies is broad and varied. Nevertheless, it is possible to extract some general trends and tendencies. In the above sections, we provided an overview to aid in understanding what cognitive functions language may and may not be involved in. In the following, we will attempt to tie it all together. Additionally, we will provide suggestions and recommendations for methodology used in future studies – in order to make results from different experiments more comparable – and encourage theoretically motivated reasons for choosing one interference type over another.

Summary of the findings

As can be seen in Table 1 and Fig. 2, it seems to be the case that verbal interference has a specific disruptive effect on tasks involving simple categorization, mental arithmetic, memory, motor control, and task switching. Verbal interference does not appear to have a specific disruptive effect on visual change, visuospatial integration and wayfinding, reasoning with non-verbal materials, or theory-of-mind processing. For the reasoning with verbal materials and complex categorization categories, the evidence appears equivocal. Generally, the studies on reasoning with verbal materials that found a specific detrimental effect of verbal interference only found this effect when participants were highly skilled or trained (Gilhooly et al., 1999; Meiser et al., 2001) or when the premises were presented sequentially (Gilhooly et al., 2002). This might suggest that participants who had learned a strategy (probably through verbal instruction) were less able to use that under verbal interference conditions, and that inner speech was used to rehearse premises continuously to keep the memory of them from degrading. The studies on complex categorization that investigated novel category learning generally demonstrate involvement of working memory, but it remains somewhat unclear whether the verbal component of working memory plays a specific role (Maddox et al., 2004; Minda et al., 2008; Newell et al., 2010; Zeithamova & Maddox, 2007). The one study that tested complex categorization by abstracting over multiple categories did find a specific effect of verbal interference (Lupyan, 2009).

When does covert language use affect task performance?

Language appears to be recruited for solving problems by cuing yourself to remember the relevant task rule, naming shades of a color to distinguish it from other colors, or naming objects or features to be remembered. There is evidence of both implicit and spontaneous language effects and more explicit language strategies – our findings suggest people sometimes use very explicit verbal strategies to solve tasks, as seen for example in the context of reasoning with verbal materials. In general, it appears that covert language

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

Table 1 Primary task areas with evidence of covert language involvement. Note that some studies used multiple interference types and thus appear more than once in the "Interference task type" and "Was there a specific effect of verbal interference?" columns

Primary task area	Number of studies included in the review	Number of participants included in the review	Interference task type (N studies)	Specific effect of verbal interference (N/total studies)	Specific effect of verbal interference (N/total participants)
Categorization (complex)	5	982	Memory (4)	2/4	224/910
			Repetition (1)	1/1	72/72
Categorization (simple)	11	702	Memory (7)	5/7	362/401
			Judgment (1)	1/1	120/120
			Repetition (3)	1/3	135/181
Mental arithmetic	10	507	Memory (5)	3/5	185/353
			Repetition (4)	4/4	130/130
			Shadowing (1)	1/1	24/24
Memory	15	2,110	Memory (2)	0/2	0/900
			Repetition (12)	10/12	918/1122
			Shadowing (1)	1/1	88/88
Motor control	2	50	Stroop task (2)	2/2	50/50
Reasoning (verbal materials)	8	900	Repetition (8)	4/8	696/900
Reasoning (non-verbal materials)	12	812	Repetition (9)	3/9	166/634
			Memory (5)	0/5	0/178
Task switching	16	1,213	Repetition (16)	16/16	1213/1213
Theory of mind	4	243	Shadowing (3)	1/3	66/196
			Memory (1)	0/1	0/47
Visual change	6	248	Shadowing (3)	2/3	101/135
			Repetition (2)	1/2	12/27
			Same/different string (1)	0/1	0/86
Visuospatial integration and wayfinding	12	1,126	Shadowing (7)	2/7	370/546
			Repetition (3)	0/3	0/364
			Word/non-word judgment (2)	0/2	0/216

aids cognition when the stimuli to be perceived, assessed, manipulated, or remembered lend themselves to a verbal code. We see this, for example, with the finding that naming objects makes them more likely to be remembered if names for their features exist, or with the finding that mental arithmetic problems demanding carry or borrow operations appear to be facilitated by language.

For categorization, the hypothesis is that covert language helps by providing a label to identify categories – this is an example of where the language effects appear to be implicit and involuntary. The fact that most of the studies reviewed indicated that verbal interference disrupts categorization fits well with the label-feedback hypothesis as proposed by Lupyan (e.g., 2012a, b). This hypothesis proposes that verbal labels – whether activated through overt or covert language use, feed-back on lower-level cognitive/perceptual processes with the effect of making them more categorical than they would be otherwise. In one study, Lupyan (2009) had participants judge which of three pictures (or words) was different from two others according either perceptual features (size, color), or more holistic thematic relationships. Under verbal interference, participants were worse at categorizing objects based on perceptual features but were still able to determine the odd one out based on thematic relationships – a pattern observed also in individuals with anomic aphasia (Cohen et al., 1980; Davidoff & Roberson, 2004; Lupyan & Mirman, 2013). Such results suggest that covert language is causally implicated in categorization tasks requiring isolation of specific dimensions (e.g., color). Recognizing that cherries and bricks, or snowmen and swans, have something in common is more difficult when language is interfered with or disrupted through a neurological insult. Additional support for this idea comes from studies using transcranial direct current stimulation (Lupyan et al., 2012; Perry & Lupyan, 2014), which have found that stimulating traditional language areas (left posterior superior temporal cortex, left inferior frontal cortex) disrupts the use of single-dimension categories.

| Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022 |

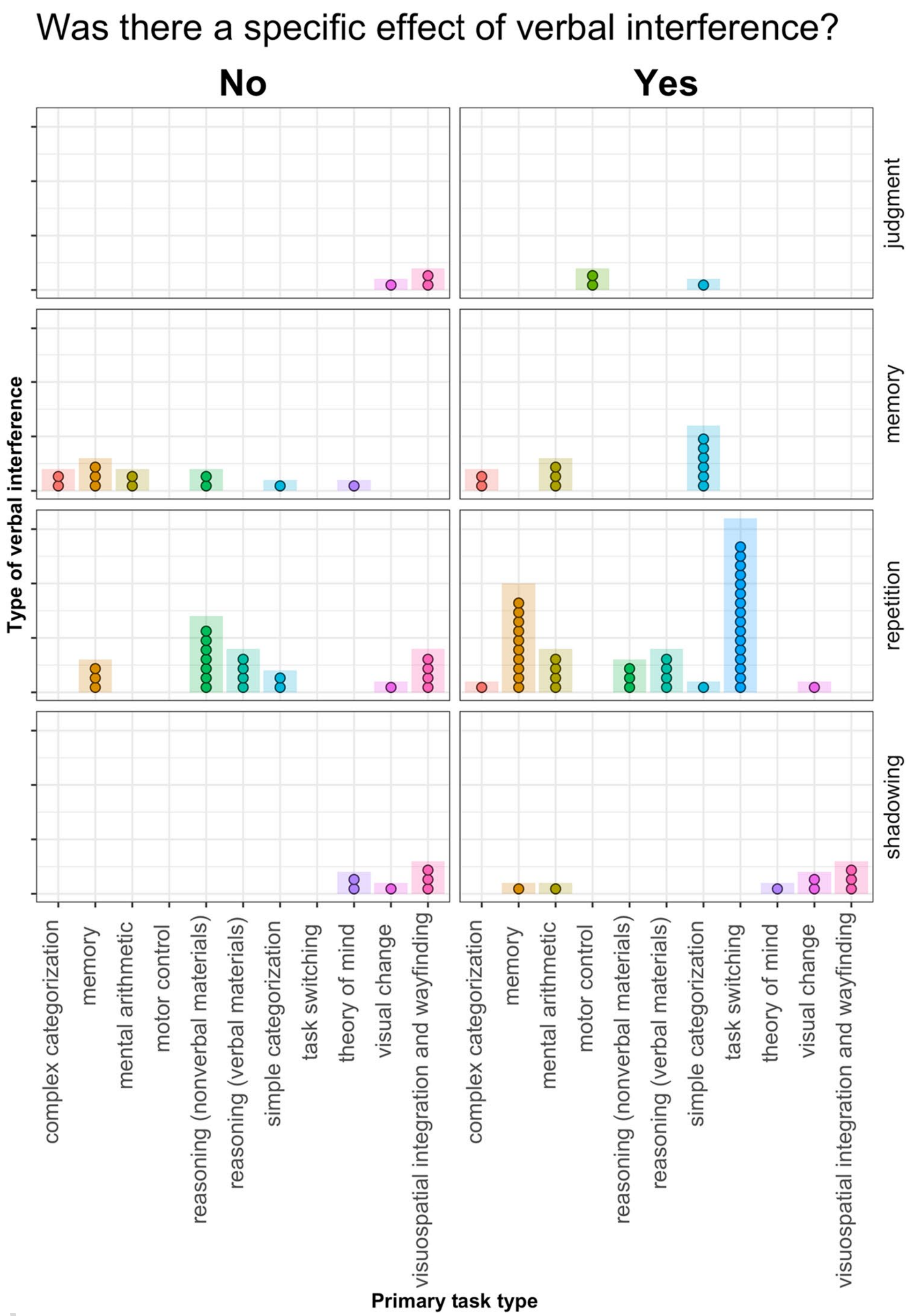

Fig. 2 Visualization of the overall results where each point represents a study included in the systematic review. The 11 primary task categories are indicated on the x axis and by color. Each row shows a different type of verbal interference. "Judgment" refers to judgment of verbal materials (for example rhyme), "memory" refers to the interference caused by a verbal memory task, "repetition" refers to repetition of simple syllables or words, and "shadowing" refers to the immediate repetition of continuously changing verbal material. Whether there was a specific effect of verbal interference (either compared with a non-verbal interference task or across different primary tasks) is indicated by the column-wise subplots in the plot grid

Aside from isolating and abstracting over specific features for categorization, language also appears to be involved in discrimination and detection of already learned categories; Roberson and Davidoff (2000) investigated recognition memory for colors and facial expressions and found that verbal interference removed the advantage normally associated

Journal : **Large 13423**	Article No : **2144**	Pages : **25**	MS Code : **PBR-TR-22-013**	Dispatch : **9-7-2022**

with categorical perception wherein cross-category judgments are more accurate than within-category judgments. Gilbert et al. (2006), (2008), Winawer et al. (2007), and He et al. (2019) all investigated color discrimination and found that there was a category advantage if the colors straddled color word boundaries and importantly that this effect disappeared with verbal interference. Roberson and Davidoff (2000) compared the effect of interference that used color words and non-color words, finding no difference between the two interference types. This suggests that the verbal interference effect they observed did not require cuing specifically task-relevant words. Interfering with language reduced categorical biases in color memory even when interference did not target color words. Converging evidence for effects of language on color memory comes from a study by Souza and Skóra (2017), who had participants remember colors while doing several tasks, among them, verbal interference and explicit color labeling (a form of up-regulation of language, see Perry & Lupyan, 2013). Unlike Roberson and Davidoff (2000), Souza and Skóra tested color memory by having participants select colors from a continuous distribution rather than through two-alternative forced choice. The authors found that explicit labeling decreased color memory in ways consistent with color labels inducing more categorical encoding in memory. Verbal interference during encoding did not affect color memory compared to control encoding conditions. A similar effect of explicit color-labeling increasing categoricality of color representations was found by Forder and Lupyan (2019), but this time on untimed color discrimination accuracy, rather than color memory.

Language does not just appear to affect cognition and perception by imposing labels and categories; however, there is also evidence that people use self-directed language to control their own behavior through rehearsal or self-cuing. In Emerson and Miyake's (2003) task-switching study, for example, verbal rehearsal plausibly helped maintain task set. This interpretation is supported by both the fact that the researchers found a specific effect of articulatory suppression and the fact that this effect depended on the existence of explicit cues to the relevant task. When there were explicit cues (plus and minus signs), articulatory suppression did not cause increased switch costs, indicating that the function of inner speech under no articulatory suppression is to provide these self-instruction cues. Asking participants to overtly verbalize the relevant cue to the task rule (presumably what they are doing covertly under normal circumstances), reduced response times, switching costs, and mixing costs (Goschke, 2000; Grange, 2013; Kirkham et al., 2012). In Nakabayashi and Burton (2008), participants were asked to remember faces – it is possible that covert language could be used as a mnemonic strategy in a similar way by allowing participants to verbalize specific features of the faces to be remembered (e.g., "potato nose," "high cheekbones," "no eyebrows", etc.) or an attempt to link faces to possible occupations or personalities. In fact, Nakabayashi and Burton (2008) found that when participants were asked to overtly describe the faces during learning, they were better at recognizing them than if they had just observed the faces silently, and Gimenes et al. (2016) found that training participants on a verbal strategy for remembering gestures improved their performance. In the four studies on reasoning with verbal materials that found specific effects of verbal interference (Farmer et al., 1986; Gilhooly et al., 1999, 2002; Meiser et al., 2001), the effects were only found for trained or highly skilled participants who had learned a specific strategy to solve the problems. As these strategies had been learned through verbal instruction, it is also likely that participants used inner speech to remind themselves of the relevant strategy for individual problems. It is also interesting that some studies found that disrupting verbal processing was associated with a loss of inhibitory control. For example, Dunbar and Sussman (1995) found that participants under verbal interference made more perseverative errors in the Wisconsin Card Sorting Task, Tullett and Inzlicht (2010) found that participants responded more impulsively on a Go/No-Go task, Wallace et al. (2017) found that participants made more excess moves on a Tower of London task while engaged in verbal interference, and both Biese et al. (2019) and Talarico et al. (2017) found that participants displayed poorer motor control while engaged in a simultaneous Stroop task.

Occasionally, effects of implicit labelling and overt strategies converge, as with nameability advantages of which there are many examples. Bek et al. (2009) investigated the Rey-Osterreith Complex Figure Test and the block design subtest of the Weschler Adult Intelligence Scale (in Experiment 1). They found that the block design task was unaffected by verbal shadowing, presumably because this task does not contain highly nameable features or require storage and rehearsal of visuospatial information. Contrastingly, copy and recall accuracy on the complex figure test were reduced if participants engaged in verbal shadowing during the copying stage and not if they were doing so during the recall stage. Verbal shadowing thus seemed to affect encoding rather than retrieval. The complex figure test notably had more nameable features than the block design test (e.g., "cross," "triangle") – participants are likely to have used these labels to support task performance and were prevented from doing so during shadowing. Further evidence for nameability advantages being sensitive to verbal interference comes from Walker and Cuthbert (1998), who investigated the unitization effect in color-shape associations. The unitization effect refers to the finding that memory for which visual properties occurred together is better if the properties are presented as belonging to the same object rather than separate objects (i.e., it is easier to remember a red triangle

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

than a triangle *and* the color red). For our present purposes, the most interesting finding of this study was that the nameability advantage for particular shapes disappeared during articulatory suppression, suggesting that some kind of verbal recoding took place under normal circumstances. In a recent related study, Zettersten and Lupyan (2020) found that more nameable features improved rule-based category learning, although they did not find that this nameability effect was modulated by verbal interference.

In summary, it appears that language can aid cognition by providing labels for better memory and faster categorization, providing self-cues for self-control, task set reminders, and verbal strategies for problem solution, and by lending a medium for rehearsal or temporary storage of items in a verbal format (as with complex mental arithmetic). Importantly, it is not only overtly verbal strategies that appear to be interrupted by verbal interference but also more involuntary or spontaneous processes. This suggests that language can influence cognition beyond the surface level.

In what kinds of tasks does covert language *not* affect performance?

The present review found little support for the on-line role of covert language in various tasks relying on primarily visual processing (the categories we named visual change, visuospatial integration and wayfinding, and reasoning using non-verbal materials). To reiterate, the hypotheses for why language would be recruited for these tasks are that language is either necessary for integrating different kinds of features (e.g., color, shape, and locations) or that visuospatial stimuli are encoded both visually and linguistically, meaning that there is somehow weaker or more shallow processing if the verbal encoding is blocked. Judging by failures to replicate the results from Hermer-Vazquez et al. (1999), however, neither the former nor the latter putative roles are strongly supported. As for the other visually based tasks, the most plausible explanation is that solving the tasks efficiently requires participants to preserve a high degree of acuity with regard to the visual stimuli (maps, complex shapes, etc.), which rarely have nameability affordances. Thus, efficient and effective processing of the stimuli does not lend itself to a verbal code, and labelling specific aspects of the stimuli is not beneficial. Interrupting verbal processing is therefore not associated with a decrement in primary task performance.

The failure to find effects of verbal interference on performance in theory-of-mind-type tasks is interesting, especially as there is a large amount of evidence supporting the idea that language and theory of mind are intimately linked in development (Astington & Baird, 2005; Astington & Jenkins, 1999; Gagne & Coppola, 2017; Lohmann & Tomasello, 2003; Milligan et al., 2007; Pyers & Senghas, 2009; Slade & Ruffman, 2005). However, there is also evidence from adults with global aphasia suggesting that their theory-of-mind abilities are intact, which means that language and theory of mind are possibly only co-dependent during development (Siegal & Varley, 2006; Varley & Siegal, 2000). As previously discussed, there are two main theories on how language facilitates theory-of-mind development: either as a representational format providing the structure for representing mental states (i.e., sentential complements) or through directing children's attention to otherwise invisible mental state dynamics. Because the present review focused on adult participants, we cannot distinguish between these two theories. These apparently conflicting findings (that language and theory of mind appear to be linked in development but not in adult cognition) can potentially be resolved either by (a) language is recruited only for development and thus ceases to be necessary once theory of mind skills are acquired, or (b) the involvement of language and theory of mind has become so automatic and proceduralized in adults that verbal interference cannot affect it.

In some interesting cases, there was a specific effect of verbal interference, but this effect was not in the direction we expected. It is important to discuss these cases as it is often assumed that if language is recruited for cognition, this will always be in a facilitative way (Dove, 2020; Dove et al., 2020). In the memory studies, for example, verbal interference in several cases caused recognition memory to decrease while actually causing mental transformation performance to *increase* (Brandimonte et al., 1992a, b; Hitch et al., 1995; Pelizzon et al., 1999). The authors of these studies interpret this as meaning that we usually encode things to be remembered verbally but that encoding in this more abstract format actually makes visual encoding less detailed and thus less available for further manipulations. In a similar vein, verbal overshadowing research indicates that forcing verbal encoding of visual stimuli can cause memory performance to deteriorate (Alogna et al., 2014; Lane & Schooler, 2004; Schooler & Engstler-Schooler, 1990). In some additional cases, verbal interference also caused primary task processing to be faster (Evans & Brooks, 1981; Forgeot d'Arc & Ramus, 2011; Phillips, 1999), perhaps indicating that converting to a verbal code under normal circumstances takes time. It is also possible that verbal interference makes participants more likely to give their initial dominant response, which can cause more errors but faster responses.

It is important to note that a null result in a verbal interference experiment does not necessarily mean that language is in no way involved with that process. It is possible that language still affects the process but off-line, as, for example, discussed with regard to theory of mind where language looks to be involved during development, but not in on-line processing in adults. It is also possible that language is involved on-line but immune to verbal interference, for instance because its involvement has become

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

so proceduralized and automatic that it can no longer be disrupted by superficial linguistic interference. This latter possibility is discussed in more detail by Wolff and Holmes (2011), who stated that "the long-term use of a language may direct habitual attention to specific properties of the world, even in nonlinguistic contexts. At a more general level, language use may also induce a given mode of processing, which may persist even as people engage in other nonlinguistic tasks … these effects of 'thinking after language' should be less attenuated by verbal interference tasks, since they occur after language is no longer in use, rather than involving the recruitment of linguistic codes during processing." (p. 259)

Choosing the interference task

It is a common problem that the different interference tasks are not matched in terms of general difficulty. One approach to this, taken by, for example, Lupyan (2009) and Hermer-Vazquez et al. (1999), is to check that the verbal and non-verbal interference tasks disrupt a third concurrent task to the same extent. This could for example be a visual search task. This approach is problematic, however, in that it glosses over the fact that the verbal and non-verbal components might also be differentially involved in this third concurrent task. It is difficult to choose a third concurrent task to validate the equivalence of the interference tasks because the literature is so divided on which tasks involve covert language and which do not. Another approach is to find a verbal and a non-verbal interference task that are in theory equivalent in every respect but their "verbality" (Perry & Lupyan, 2013), including performance. This approach faces challenges because tasks that are equivalent in everything but their verbality may yet place different demands on attention and executive function. Ideally, the tasks should at least be equated as separate single tasks in terms of their difficulty, and performance should neither be at ceiling nor at floor. This would make it possible to analyze potential trade-off effects with the primary task.

As we have seen, there are four types of verbal interference that have been used: syllable repetition, verbal memory, verbal shadowing, and judgment tasks. Only too rarely have the different interference tasks been directly compared, even though they might yield different predictions depending on which aspect of language (rehearsal, syntactic structure, verbal labels) you hypothesize is involved in the primary task you are investigating. Bek et al. (2009, 2013) directly compared syllable shadowing and prose shadowing, which should intuitively be different in terms of which components of language are involved. After all, syllable repetition uses less "language" than prose shadowing (semantics, syntax, morphology, etc.), which is precisely why syllable repetition is so widely used in working memory studies. In these experiments, there was no difference between shadowing syllables and shadowing prose. If anything, shadowing syllables resulted in a marginally more detrimental effect on visuospatial reorientation. A possible explanation may be that syllable shadowing lacks the predictability of prose shadowing and thus actually requires more cognitive resources.

Current forms of verbal interference (see above) are not well suited for distinguishing which components of language are most involved in performance on the primary task. Comparing interference involving task-relevant versus task-irrelevant words (Piccardi et al., 2020; Roberson & Davidoff, 2000) offers some, albeit limited, insights. A promising avenue for future research would be to compare manipulations designed to increase language involvement (e.g., as in Forder & Lupyan, 2019; Lupyan, 2008; Lupyan & Swingley, 2012) with conditions suppressing language involvement (e.g., as was done by Souza & Skóra, 2017). Once verbal interference has indicated that language in some form may be involved, up-regulating language involvement would be better suited to targeting specific hypotheses about components of language involved. We see this for example in findings indicating that the way language helps task switching is by helping to cue the relevant task rule (Goschke, 2000; Grange, 2013; Kirkham et al., 2012). Without additional task manipulations supplementing the dual-task interference, we would not have much indication as to *how* language helps task switching performance. Another example of up-regulating language shedding light on the specific ways language may be involved comes from the sport psychology literature where self-talk interventions (up-regulating language) are much more common than dual-task interference studies (Hatzigeorgiadis et al., 2011; Tod et al., 2011). Here, participants are often trained to use different types of self-directed verbalizations (instructional vs. motivational, positive vs. negative, etc.), which result in different effects on performance depending on the participant's skill level (Zourbanos et al., 2013), the motor demands of the sport (Theodorakis et al., 2000), and whether the self-talk takes place in a competition or practice context (Hatzigeorgiadis et al., 2014). In addition to focusing on the content of internal verbalizations, it is also important to understand the *stage* at which interfering with language affects performance, for example, during memory encoding, retrieval, or both (Frank et al., 2012; Nakabayashi & Burton, 2008). This may help tease apart effects of verbal encoding (nameability effects in memory, verbal overshadowing) and "mental workspace" functions (using the phonological loop to keep track of carry or borrow operations, keeping track of the relevant task rule). Future studies would benefit from clarifying their predictions about language involvement in this way.

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

Summary of suggestions for future studies

Future studies should follow these recommendations:

1. Include control conditions of both the primary and the secondary tasks.
2. Make theoretically informed and hypothesis-driven choices about the type of interference task and/or directly compare effects of different types.
3. Ensure that the different interference tasks are matched in terms of difficulty/attentional demands by measuring performance.
4. Consider potential trade-offs between effort/resources put into the primary tasks and the secondary tasks.
5. Delineate the precise mechanisms by which language is expected to help cognition.

Conclusion

It appears that language – including inner speech – is a powerful tool for directing attention, improving memory, and controlling actions. These three processes, however, are intimately connected. For example, paying attention to specific aspects or properties of something makes it more likely that you will remember it later, and remembering how you acted in a past situations can (and should) influence what you attend to and how you act in the current situation. We reviewed 101 studies investigating the on-line role of language in some cognitive function using a dual-task interference methodology. Overall, we found that it is likely the case that covert language is recruited for behavioral self-cuing (inhibitory control, task set reminders, verbal strategy), rehearsal for memory when items to be remembered have readily available labels, and as a workspace for complex mental arithmetic. We found less evidence for a role of on-line language use in cross-modal integration, reasoning that relies on a high degree of visual detail (such as map tasks, visual recursion tasks, and some matrix problems), and theory of mind. It is important to note that we only examined *one* way of investigating the role of language in cognition and that other patterns of effects may appear with the use of different approaches. Interestingly, we found that recruiting language for non-verbal tasks is not always purely advantageous, but may present costs in term of processing speed, loss of visual detail, and verbal overshadowing. Future studies should include relevant control conditions for both primary and secondary tasks, make informed and justified decisions about the interference tasks, ensure that the interference tasks are appropriately matched, and delineate the precise mechanisms by which covert language is expected to help cognition in the on-line processing of a given primary task.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13423-022-02144-7.

Declarations

This work was supported by a PhD grant awarded to the first author. The authors all declare no conflicts of interest. The full table of the papers included in this review can be found in Appendix A, which is available as Online Supplementary Material.

References

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin, 141*(5), 931–965. https://doi.org/10.1037/bul0000021

Alogna, V. K., Attaya, M. K., Aucoin, P., Bahník, Š., Birch, S., Birt, A. R., Bornstein, B. H., Bouwmeester, S., Brandimonte, M. A., Brown, C., Buswell, K., Carlson, C., Carlson, M., Chu, S., Cislak, A., Colarusso, M., Colloff, M. F., Dellapaolera, K. S., Delvenne, J.-F., … Zwaan, R. A. (2014). Registered Replication Report: Schooler and Engstler-Schooler (1990). *Perspectives on Psychological Science*, *9*(5), 556–578. https://doi.org/10.1177/1745691614545653

Annett, J. M., & Leslie, J. C. (1996). Effects of visual and verbal interference tasks on olfactory memory: The role of task complexity. *British Journal of Psychology, 87*(3), 447–460. https://doi.org/10.1111/j.2044-8295.1996.tb02601.x

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition, 1*(1), 3–34.

Astington, J. W., & Baird, J. A. (2005). *Why language matters for theory of mind*. Oxford University Press.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*(5), 1311.

Baddeley, A. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18*(4), 362–365. https://doi.org/10.1080/14640746608400055

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47–89). Elsevier. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A., & Larsen, J. D. (2007). The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative. *Quarterly Journal of Experimental Psychology, 60*(4), 497–504. https://doi.org/10.1080/17470210601147572

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General, 130*(4), 641–657. https://doi.org/10.1037/0096-3445.130.4.641

Baldo, J., Dronkers, N., Wilkins, D., Ludy, C., Raskin, P., & Kim, J. (2005). Is problem solving dependent on language? *Brain and Language, 92*(3), 240–250. https://doi.org/10.1016/j.bandl.2004.06.103

Bek, J., Blades, M., Siegal, M., & Varley, R. (2009). Linguistic processes in visuospatial representation: Clarifying verbal interference effects. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *31*.

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

Bek, J., Blades, M., Siegal, M., & Varley, R. (2013). Dual-task interference in spatial reorientation: Linguistic and nonlinguistic factors. *Spatial Cognition & Computation, 13*(1), 26–49. https://doi.org/10.1080/13875868.2011.590622

Bermúdez, J. L. (2003). *Thinking without words*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159691.001.0001

Biese, K. M., Pietrosimone, L. E., Andrejchak, M., Lynall, R. C., Wikstrom, E. A., & Padua, D. A. (2019). Preliminary investigation on the effect of cognition on jump-landing performance using a clinically relevant setup. *Measurement in Physical Education and Exercise Science, 23*(1), 78–88. https://doi.org/10.1080/1091367X.2018.1518875

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. (1992). Influence of short-term memory codes on visual image processing: Evidence from image transformation tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(1), 157–165. https://doi.org/10.1037/0278-7393.18.1.157

Brandimonte, M. A., Hitch, G. J., & Bishop, D. V. M. (1992). Verbal recoding of visual stimuli impairs mentalimagetransformations. *Memory & Cognition, 20*(4), 449–455. https://doi.org/10.3758/BF03210929

Brown, R. G., & Marsden, C. D. (1991). Dual task performance and processing resources in normal subjects and patients with Parkinson's disease. *Brain, 114A*(1), 215–231. https://doi.org/10.1093/oxfordjournals.brain.a101858

Bryck, R. L., & Mayr, U. (2005). On the role of verbalization during task set selection: Switching or serial order control? *Memory & Cognition, 33*(4), 611–623. https://doi.org/10.3758/BF03195328

Caffò, A. O., Picucci, L., Di Masi, M. N., & Bosco, A. (2011). Working memory components and virtual reorientation: A dual-task study. In E. S. Levin (Ed.), *Working memory: Capacity, developments and improvement techniques* (pp. 249–266). Nova Science Publisher.

Camos, V., Mora, G., & Barrouillet, P. (2013). Phonological similarity effect in complex span task. *Quarterly Journal of Experimental Psychology, 66*(10), 1927–1950. https://doi.org/10.1080/17470218.2013.768275

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences, 25*(6), 657–674. https://doi.org/10.1017/S0140525X02000122

Cheetham, J. M., Rahm, B., Kaller, C. P., & Unterrainer, J. M. (2012). Visuospatial over verbal demands in predicting Tower of London planning tasks: Visuospatial demands during TOL planning tasks. *British Journal of Psychology, 103*(1), 98–116. https://doi.org/10.1111/j.2044-8295.2011.02049.x

Clark, A. (1998). Magic words: How language augments human computation. In P. Carruthers & J. Boucher (Eds.), *Language and thought* (1st ed., pp. 162–183). Cambridge University Press. https://doi.org/10.1017/CBO9780511597909.011

Clearman, J., Klinger, V., & Szűcs, D. (2017). Visuospatial and verbal memory in mental arithmetic. *Quarterly Journal of Experimental Psychology, 70*(9), 1837–1855. https://doi.org/10.1080/17470218.2016.1209534

Coetzee, J., Monti, M., Iacoboni, M., Wu, A., & Johnson, M. (2019). Separability of logic and language: A TMS study. *Brain Stimulation, 12*(2), 543. https://doi.org/10.1016/j.brs.2018.12.792

Cohen, R., Kelter, S., & Woll, G. (1980). Analytical competence and language impairment in aphasia. *Brain and Language, 10*(2), 331–347. https://doi.org/10.1016/0093-934X(80)90060-7

Conrad, R. (1964). Acoustic confusion in immediate memory. *British Journal of Psychology, 55*, 75–84.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology, 55*, 429–432.

Croijmans, I., Arshamian, A., Speed, L. J., & Majid, A. (2021). Wine experts' recognition of wine odors is not verbally mediated. *Journal of Experimental Psychology: General, 150*(3), 545–559. https://doi.org/10.1037/xge0000949

Davidoff, J., & Roberson, D. (2004). Preserved thematic and impaired taxonomic categorisation: A case study. *Language and Cognitive Processes, 19*(1), 137–174. https://doi.org/10.1080/01690960344000125

De Renzi, E., & Spinnler, H. (1967). Impaired performance on color tasks in patients with hemispheric damage. *Cortex, 3*(2), 194–217. https://doi.org/10.1016/S0010-9452(67)80012-1

de Villiers, J. (2007). The interface of language and theory of mind. *Lingua, 117*(11), 1858–1878.

de Villiers, J., & de Villiers, P. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191–228). Psychology Press.

de Villiers, J., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17*(1), 1037–1060.

Dent, K. (2009). Coding categorical and coordinate spatial relations in visual–spatial short-term memory. *Quarterly Journal of Experimental Psychology, 62*(12), 2372–2387. https://doi.org/10.1080/17470210902853548

Dessalegn, B., & Landau, B. (2008). More than meets the eye: The role of language in binding and maintaining feature conjunctions. *Psychological Science, 19*(2), 189–195. https://doi.org/10.1111/j.1467-9280.2008.02066.x

DeStefano, D., & LeFevre, J. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*(3), 353–386. https://doi.org/10.1080/09541440244000328

Dove, G. (2020). More than a scaffold: Language is a neuroenhancement. *Cognitive Neuropsychology, 37*(5–6), 288–311.

Dove, G., Barca, L., Tummolini, L., & Borghi, A. M. (2020). Words have a weight: Language as a source of inner grounding and flexibility in abstract concepts. *Psychological Research*, 1–17.

Dunbar, R. I. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews, 6*(5), 178–190.

Dunbar, K., & Sussman, D. (1995). Toward a cognitive account of frontal lobe function: Simulating frontal lobe deficits in normal subjects. *Annals of the New York Academy of Sciences, 769*(1 Structure and), 289–304. https://doi.org/10.1111/j.1749-6632.1995.tb38146.x

Dungan, J., & Saxe, R. (2012). Matched false-belief performance during verbal and nonverbal interference. *Cognitive Science, 36*(6), 1148–1156. https://doi.org/10.1111/j.1551-6709.2012.01248.x

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language, 48*(1), 148–168. https://doi.org/10.1016/S0749-596X(02)00511-9

Evans, J. S. B. T., & Brooks, P. G. (1981). Competing with reasoning: A test of the working memory hypothesis. *Current Psychology, 1*(2), 139–147. https://doi.org/10.1007/BF02684486

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a visuo-spatial scratch-pad in working memory. *The Quarterly Journal of Experimental Psychology Section A, 38*(4), 675–688. https://doi.org/10.1080/14640748608401620

Feinmann, D. (2020). Language and thought in the motion domain: Methodological considerations and new empirical evidence. *Journal of Psycholinguistic Research, 49*(1), 1–29. https://doi.org/10.1007/s10936-019-09668-5

Fodor, J. A. (1975). *The language of thought*. Harvard University Press.

Forder, L., & Lupyan, G. (2019). Hearing words changes color perception: Facilitation of color discrimination by verbal and visual cues. *Journal of Experimental Psychology: General, 148*(7), 1105–1123. https://doi.org/10.1037/xge0000560

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

Forgeot d'Arc, B., & Ramus, F. (2011). Belief attribution despite verbal interference. *Quarterly Journal of Experimental Psychology, 64*(5), 975–990. https://doi.org/10.1080/17470218.2010.524413

Frank, M. C., Fedorenko, E., Lai, P., Saxe, R., & Gibson, E. (2012). Verbal interference suppresses exact numerical representation. *Cognitive Psychology, 64*(1–2), 74–92. https://doi.org/10.1016/j.cogpsych.2011.10.004

Gagne, D. L., & Coppola, M. (2017). Visible social interactions do not support the development of false belief understanding in the absence of linguistic input: Evidence from deaf adult homesigners. *Frontiers in Psychology, 8,* 837. https://doi.org/10.3389/fpsyg.2017.00837

Gaillard, V., Destrebecqz, A., & Cleeremans, A. (2012). The influence of articulatory suppression on the control of implicit sequence knowledge. *Frontiers in Human Neuroscience, 6*. https://doi.org/10.3389/fnhum.2012.00208

Garden, S., Cornoldi, C., & Logie, R. H. (2002). Visuo-spatial working memory in navigation. *Applied Cognitive Psychology, 16*(1), 35–50. https://doi.org/10.1002/acp.746

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences, 103*(2), 489–494. https://doi.org/10.1073/pnas.0509868103

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2008). Support for lateralization of the Whorf effect beyond the realm of color discrimination. *Brain and Language, 105*(2), 91–98. https://doi.org/10.1016/j.bandl.2007.06.001

Gilhooly, K. J. (2005). Working memory and strategies in reasoning. In E. Newton & M. Roberts (Eds.), *Methods of thought: Individual differences in reasoning strategies* (1st ed., pp. 57–80).

Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory & Cognition, 21*(1), 115–124. https://doi.org/10.3758/BF03211170

Gilhooly, K. J., Logie, R. H., & Wynn, V. (1999). Syllogistic reasoning tasks, working memory, and skill. *European Journal of Cognitive Psychology, 11*(4), 473–498. https://doi.org/10.1080/095414499382264

Gilhooly, K. J., Wynn, V., Phillips, L. H., Logie, R. H., & Sala, S. D. (2002). Visuo-spatial and verbal working memory in the five-disc Tower of London task: An individual differences approach. *Thinking & Reasoning, 8*(3), 165–178. https://doi.org/10.1080/13546780244000006

Gimenes, G., Pennequin, V., & Mercer, T. (2016). What is the best strategy for retaining gestures in working memory? *Memory, 24*(6), 757–765. https://doi.org/10.1080/09658211.2015.1049544

Gleitman, L., & Papafragou, A. (2005). Language and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 633–661). Cambridge University Press.

Gomila, A., Travieso, D., & Lobo, L. (2012). Wherein is human cognition systematic? *Minds and Machines, 22*(2), 101–115. https://doi.org/10.1007/s11023-012-9277-z

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In S. Monsell & J. Driver (Eds.), *Attention & performance XVIII: Control of cognitive processes*. MIT Press.

Grange, J. A. (2013). Manipulating task preparation via verbalisations does not affect inhibition in set switching. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2350387

Hatzigeorgiadis, A., Zourbanos, N., Galanis, E., & Theodorakis, Y. (2011). Self-talk and sports performance: A meta-analysis. *Perspectives on Psychological Science, 6*(4), 348–356. https://doi.org/10.1177/1745691611413136

Hatzigeorgiadis, A., Galanis, E., Zourbanos, N., & Theodorakis, Y. (2014). Self-talk and Competitive Sport Performance. *Journal of Applied Sport Psychology, 26*(1), 82–95. https://doi.org/10.1080/10413200.2013.790095

He, H., Li, J., Xiao, Q., Jiang, S., Yang, Y., & Zhi, S. (2019). Language and color perception: Evidence from Mongolian and Chinese speakers. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.00551

Hegarty, M., Shah, P., & Miyake, A. (2000). Constraints on using the dual-task methodology to specify the degree of central executive involvement in cognitive tasks. *Memory & Cognition, 28*(3), 376–385. https://doi.org/10.3758/BF03198553

Henson, R., Hartley, T., Burgess, N., Hitch, G., & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing-signal hypothesis. *The Quarterly Journal of Experimental Psychology Section A, 56*(8), 1307–1334. https://doi.org/10.1080/02724980244000747

Hermer-Vazquez, L., Spelke, E. S., & Katsnelson, A. S. (1999). Sources of flexibility in human cognition: dual-task studies of space and language. *Cognitive Psychology, 39*(1), 3–36. https://doi.org/10.1006/cogp.1998.0713

Hintzman, D. L. (1967). Articulatory coding in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 6*(3), 312–316.

Hitch, G. J., Brandimonte, M. A., & Walker, P. (1995). Two types of representation in visual memory: Evidence from the effects of stimulus contrast on image combination. *Memory & Cognition, 23*(2), 147–154. https://doi.org/10.3758/BF03197217

Hollingworth, A. (2003). Failures of retrieval and comparison constrain change detection in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 388–403. https://doi.org/10.1037/0096-1523.29.2.388

Hund, A. M. (2016). Visuospatial working memory facilitates indoor wayfinding and direction giving. *Journal of Environmental Psychology, 45,* 233–238. https://doi.org/10.1016/j.jenvp.2016.01.008

Hupbach, A., Hardt, O., Nadel, L., & Bohbot, V. D. (2007). Spatial reorientation: Effects of verbal and spatial shadowing. *Spatial Cognition & Computation, 7*(2), 213–226. https://doi.org/10.1080/13875860701418206

Imbo, I., & LeFevre, J.-A. (2010). The role of phonological and visual working memory in complex arithmetic for Chinese- and Canadian-educated adults. *Memory & Cognition, 38*(2), 176–185. https://doi.org/10.3758/MC.38.2.176

Imbo, I., De Rammelaere, S., & Vandierendonck, A. (2005). New insights in the role of working memory in carry and borrow operations. *Psychologica Belgica, 45*(2), 101–121. https://doi.org/10.5334/pb-45-2-101

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(3), 656–674.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2007). The phonological store abandoned. *Quarterly Journal of Experimental Psychology, 60*(4), 505–511. https://doi.org/10.1080/17470210601147598

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin, 136*(5), 849–874. https://doi.org/10.1037/a0019842

Kirkham, A. J., Breeze, J. M., & Marí-Beffa, P. (2012). The impact of verbal instructions on goal-directed behaviour. *Acta Psychologica, 139*(1), 212–219. https://doi.org/10.1016/j.actpsy.2011.09.016

Klauer, K. C. (1997). Working memory involvement in propositional and spatial reasoning. *Thinking & Reasoning, 3*(1), 9–47. https://doi.org/10.1080/135467897394419

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking—An integrative review of dual-task and task-switching research. *Psychological Bulletin, 144*(6), 557–583. https://doi.org/10.1037/bul0000144

Kranjec, A., Lupyan, G., & Chatterjee, A. (2014). Categorical Biases in Perceiving Spatial Relations. *PLoS ONE, 9*(5), e98604. https://doi.org/10.1371/journal.pone.0098604

Labate, E., Pazzaglia, F., & Hegarty, M. (2014). What working memory subcomponents are needed in the acquisition of survey knowledge? Evidence from direction estimation and shortcut tasks. *Journal of Environmental Psychology, 37,* 73–79. https://doi.org/10.1016/j.jenvp.2013.11.007

Lane, S. M., & Schooler, J. W. (2004). Skimming the surface: Verbal overshadowing of analogical retrieval. *Psychological Science, 15*(11), 715–719. https://doi.org/10.1111/j.0956-7976.2004.00747.x

Law, A. S., Logie, R. H., & Pearson, D. G. (2006). The impact of secondary tasks on multitasking in a virtual environment. *Acta Psychologica, 122*(1), 27–44. https://doi.org/10.1016/j.actpsy.2005.09.002

Law, A. S., Trawley, S. L., Brown, L. A., Stephens, A. N., & Logie, R. H. (2013). The impact of working memory load on task execution and online plan adjustment during multitasking in a virtual environment. *Quarterly Journal of Experimental Psychology, 66*(6), 1241–1258. https://doi.org/10.1080/17470218.2012.748813

Lee, K.-M., & Kang, S.-Y. (2002). Arithmetic operation and working memory: Differential suppression in dual tasks. *Cognition, 83*(3), B63–B68. https://doi.org/10.1016/S0010-0277(02)00010-0

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology, 58*(4), 487–524. https://doi.org/10.1016/j.cogpsych.2008.12.001

Liefooghe, B., Vandierendonck, A., Muyllaert, I., Verbruggen, F., & Vanneste, W. (2005). The phonological loop in task alternation and task repetition. *Memory, 13*(5), 550–560. https://doi.org/10.1080/09658210444000250

Liu, Q., Chen, A.-T., Wang, Q., Zhou, L., & Sun, H.-J. (2008). An evidence for the effect of categorical perception on color perception: An evidence for the effect of categorical perception on color perception. *Acta Psychologica Sinica, 40*(1), 8–13. https://doi.org/10.3724/SP.J.1041.2008.00008

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition, 22*(4), 395–410. https://doi.org/10.3758/BF03200866

Lohmann, H., & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development, 74*(4), 1130–1144. https://doi.org/10.1111/1467-8624.00597

Luo, X., Sexton, N. J., & Love, B. C. (2021). A deep learning account of how language affects thought. *Language, Cognition and Neuroscience*, 1–10. https://doi.org/10.1080/23273798.2021.2001023

Lupyan, G. (2008). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition, 108*(2), 566–577.

Lupyan, G. (2009). Extracommunicative functions of language: Verbal interference causes selective categorization impairments. *Psychonomic Bulletin & Review, 16*(4), 711–718. https://doi.org/10.3758/PBR.16.4.711

Lupyan, G. (2012a). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology, 3.* https://doi.org/10.3389/fpsyg.2012.00054

Lupyan, G. (2012b). What do words do? Towards a theory of language-augmented thought. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 57, pp. 255–297). Academic Press. http://www.sciencedirect.com/science/article/pii/B9780123942937000078

Lupyan, G., & Mirman, D. (2013). Linking language and categorization: Evidence from aphasia. *Cortex, 49*(5), 1187–1194. https://doi.org/10.1016/j.cortex.2012.06.006

Lupyan, G., & Swingley, D. (2012). Self-directed speech affects visual search performance. *The Quarterly Journal of Experimental Psychology, 65*(6), 1068–1085. https://doi.org/10.1080/17470218.2011.647039

Lupyan, G., Mirman, D., Hamilton, R., & Thompson-Schill, S. L. (2012). Categorization is modulated by transcranial direct current stimulation over left prefrontal cortex. *Cognition, 124*(1), 36–49. https://doi.org/10.1016/j.cognition.2012.04.002

Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes, 66*(3), 309–332. https://doi.org/10.1016/j.beproc.2004.03.011

Maddox, W. T., Ashby, F. G., Ing, A. D., & Pickering, A. D. (2004). Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory & Cognition, 32*(4), 582–591. https://doi.org/10.3758/BF03195849

Malle, B. F. (2002). The relation between language and theory of mind in development and evolution. In T. Givón & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265–284). John Benjamins Publishing Company.

Martins, M. de J. D., Muršič, Z., Oh, J., & Fitch, W. T. (2015). Representing visual recursion does not require verbal or motor resources. *Cognitive Psychology, 77,* 20–41. https://doi.org/10.1016/j.cogpsych.2015.01.004

Meilinger, T., Knauff, M., & Bülthoff, H. H. (2008). Working memory in wayfinding-a dual task experiment in a virtual City. *Cognitive Science, 32*(4), 755–770. https://doi.org/10.1080/03640210802067004

Meiser, T., Klauer, K. C., & Naumer, B. (2001). Propositional reasoning and working memory: The role of prior training and pragmatic content q. *Acta Psychologica, 25*.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78*(2), 622–646.

Minda, J. P., Desroches, A. S., & Church, B. A. (2008). Learning rule-described and non-rule-described categories: A comparison of children and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1518–1533. https://doi.org/10.1037/a0013355

Mitsuhashi, S., Hirata, S., & Okuzumi, H. (2018). Role of inner speech on the Luria hand test. *Cogent Psychology, 5*(1), 1449485. https://doi.org/10.1080/23311908.2018.1449485

Miyake, A., Witzki, A. H., & Emerson, M. J. (2001). Field dependence–independence from a working memory perspective: A dual-task investigation of the Hidden Figures Test. *Memory, 9*(4–6), 445–457. https://doi.org/10.1080/09658210143000029

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*(2–3), 123–142. https://doi.org/10.1016/j.actpsy.2003.12.004

Müller, M. (1978). The science of thought. In *The science of thought* / (Repr.). AMS Press.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 21*(3), 263–276.

Nakabayashi, K., & Burton, A. M. (2008). The role of verbal processing at different stages of recognition memory for faces. *European Journal of Cognitive Psychology, 20*(3), 478–496. https://doi.org/10.1080/09541440801946174

Journal : Large 13423	Article No : 2144	Pages : 25	MS Code : PBR-TR-22-013	Dispatch : 9-7-2022

Newell, B. R., Dunn, J. C., & Kalish, M. (2010). The dimensionality of perceptual category learning: A state-trace analysis. *Memory & Cognition, 38*(5), 563–581. https://doi.org/10.3758/MC.38.5.563

Newton, A. M., & de Villiers, J. G. (2007). Thinking while talking: Adults fail nonverbal false-belief reasoning. *Psychological Science, 18*(7), 574–579. https://doi.org/10.1111/j.1467-9280.2007.01942.x

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 45*(3), 255–287. https://doi.org/10.1037/h0084295

Pelizzon, L., Brandimonte, M. A., & Favretto, A. (1999). Imagery and recognition: Dissociable measures of memory? *European Journal of Cognitive Psychology, 11*(3), 429–443. https://doi.org/10.1080/713752323

Perkins, J., & McLaughlin Cook, N. (1990). Recognition and recall of odours: The effects of suppressing visual and verbal encoding processes. *British Journal of Psychology, 81*(2), 221–226. https://doi.org/10.1111/j.2044-8295.1990.tb02357.x

Perry, L. K., & Lupyan, G. (2013). What the online manipulation of linguistic activity can tell us about language and thought. *Frontiers in Behavioral Neuroscience, 7*. https://doi.org/10.3389/fnbeh.2013.00122

Perry, L. K., & Lupyan, G. (2014). The role of language in multidimensional categorization: Evidence from transcranial direct current stimulation and exposure to verbal labels. *Brain and Language, 135,* 66–72. https://doi.org/10.1016/j.bandl.2014.05.005

Peterson, L. R. (1969). Concurrent verbal activity. *Psychological Review, 76*(4), 376–386.

Phillips, L. H. (1999). The role of memory in the tower of London task. *Memory, 7*(2), 209–231. https://doi.org/10.1080/741944066

Phillips, L. H., Wynn, V., Gilhooly, K. J., Della Sala, S., & Logie, R. H. (1999). The role of memory in the tower of London task. *Memory, 7*(2), 209–231. https://doi.org/10.1080/741944066

Piccardi, L., Bocchi, A., Palmiero, M., Boccia, M., D'Amico, S., & Nori, R. (2020). Chatting while walking does not interfere with topographical working memory. *Brain Sciences, 10*(11), 811. https://doi.org/10.3390/brainsci10110811

Pilling, M., Wiggett, A., Özgen, E., & Davies, I. R. L. (2003). Is color "categorical perception" really perceptual? *Memory & Cognition, 31*(4), 538–551. https://doi.org/10.3758/BF03196095

Pyers, J. E., & Senghas, A. (2009). Language promotes false-belief understanding: evidence from learners of a new sign language. *Psychological Science, 20*(7), 805–812. https://doi.org/10.1111/j.1467-9280.2009.02377.x

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*(2), 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Rao, K. V., & Baddeley, A. (2013). Raven's matrices and working memory: A dual-task approach. *Quarterly Journal of Experimental Psychology, 66*(10), 1881–1887. https://doi.org/10.1080/17470218.2013.828314

Ratliff, K. R., & Newcombe, N. S. (2005). *Human spatial reorientation using dual task paradigms* (pp. 6).

Ratliff, K. R., & Newcombe, N. S. (2008). Is language necessary for human spatial reorientation? Reconsidering evidence from dual task paradigms. *Cognitive Psychology, 56*(2), 142–163. https://doi.org/10.1016/j.cogpsych.2007.06.002

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: The effect of verbal interference. *Memory & Cognition, 28*(6), 977–986. https://doi.org/10.3758/BF03209345

Robert, N. D., & LeFevre, J.-A. (2013). Ending up with less: The role of working memory in solving simple subtraction problems with positive and negative answers. *Research in Mathematics Education, 15*(2), 165–176. https://doi.org/10.1080/14794802.2013.797748

Romano, S., Salles, A., Amalric, M., Dehaene, S., Sigman, M., & Figueira, S. (2018). Bayesian validation of grammar productions for the language of thought. *PLOS ONE, 13*(7), e0200420. https://doi.org/10.1371/journal.pone.0200420

Saeki, E. (2007). Phonological loop and goal maintenance: effect of articulatory suppression in number-size consistency task. *Psychologia, 50*(2), 122–131. https://doi.org/10.2117/psysoc.2007.122

Saeki, E., & Saito, S. (2004a). The role of the phonological loop in task switching performance: The effect of articulatory suppression in the alternating runs paradigm. *Psychologia, 47*(1), 35–43. https://doi.org/10.2117/psysoc.2004.35

Saeki, E., & Saito, S. (2004b). Effect of articulatory suppression on task-switching performance: Implications for models of working memory. *Memory, 12*(3), 257–271. https://doi.org/10.1080/09658210244000649

Saeki, E., & Saito, S. (2009). Verbal representation in task order control: An examination with transition and task cues in random task switching. *Memory & Cognition, 37*(7), 1040–1050. https://doi.org/10.3758/MC.37.7.1040

Saeki, E., Saito, S., & Kawaguchi, J. (2006). Effects of response–stimulus interval manipulation and articulatory suppression on task switching. *Memory, 14*(8), 965–976. https://doi.org/10.1080/09658210601008973

Saeki, E., Baddeley, A., Hitch, G. J., & Saito, S. (2013). Breaking a habit: A further role of the phonological loop in action control. *Memory & Cognition, 41*(7), 1065–1078. https://doi.org/10.3758/s13421-013-0320-y

Samuel, S., Durdevic, K., Legg, E. W., Lurz, R., & Clayton, N. S. (2019). Is language required to represent others' mental states? Evidence from beliefs and other representations. *Cognitive Science, 43*(1), e12710. https://doi.org/10.1111/cogs.12710

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22*(1), 36–71. https://doi.org/10.1016/0010-0285(90)90003-M

Seitz, K., & Schumann-Hengsteler, R. (2000). Mental multiplication and working memory. *European Journal of Cognitive Psychology, 12*(4), 552–570. https://doi.org/10.1080/095414400750050231

Seitz, K., & Schumann-Hengsteler, R. (2002). Phonological loop and central executive processes in mental addition and multiplication. *Psychologische Beiträge, 44*(2), 275.

Sense, F., Morey, C. C., Prince, M., Heathcote, A., & Morey, R. D. (2017). Opportunity for verbalization does not improve visual change detection performance: A state-trace analysis. *Behavior Research Methods, 49*(3), 853–862. https://doi.org/10.3758/s13428-016-0741-1

Siegal, M., & Varley, R. (2006). Aphasia, language, and theory of mind. *Social Neuroscience, 1*(3–4), 167–174.

Simons, D. J. (1996). In sight, out of mind: When object representations fail. *Psychological Science, 7*(5), 301–305. https://doi.org/10.1111/j.1467-9280.1996.tb00378.x

Sims, V. K., & Hegarty, M. (1997). Mental animation in the visuospatial sketchpad: Evidence from dual-task studies. *Memory & Cognition, 25*(3), 321–332. https://doi.org/10.3758/BF03211288

Slade, L., & Ruffman, T. (2005). How language does (and does not) relate to theory of mind: A longitudinal study of syntax, semantics, working memory and false belief. *British Journal of Developmental Psychology, 23*(1), 117–141. https://doi.org/10.1348/026151004X21332

Sokolov, A. N. (1968). *Inner speech and thought* (G. T. Onischenko, Trans.). Plenum Press.

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

Souza, A. S., & Skóra, Z. (2017). The interplay of language and visual perception in working memory. *Cognition, 166,* 277–297. https://doi.org/10.1016/j.cognition.2017.05.038

Talarico, M. K., Lynall, R. C., Mauntel, T. C., Weinhold, P. S., Padua, D. A., & Mihalik, J. P. (2017). Static and dynamic single leg postural control performance during dual-task paradigms. *Journal of Sports Sciences, 35*(11), 1118–1124. https://doi.org/10.1080/02640414.2016.1211307

Theodorakis, Y., Weinberg, R., Natsis, P., Douma, I., & Kazakas, P. (2000). The effects of motivational versus instructional self-talk on improving motor performance. *The Sport Psychologist, 14*(3), 253–271.

Tod, D., Hardy, J., & Oliver, E. (2011). Effects of self-talk: A systematic review. *Journal of Sport and Exercise Psychology, 33*(5), 666–687. https://doi.org/10.1123/jsep.33.5.666

Toms, M., Morris, N., & Ward, D. (1993). Working memory and conditional reasoning. *The Quarterly Journal of Experimental Psychology Section A, 46*(4), 679–699. https://doi.org/10.1080/14640749308401033

Trbovich, P. L., & LeFevre, J.-A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition, 31*(5), 738–745. https://doi.org/10.3758/BF03196112

Tullett, A. M., & Inzlicht, M. (2010). The voice of self-control: Blocking the inner voice increases impulsive responding. *Acta Psychologica, 135*(2), 252–256. https://doi.org/10.1016/j.actpsy.2010.07.008

van der Ham, I. J. M., & Borst, G. (2011). The nature of categorical and coordinate spatial relation processing: An interference study. *Journal of Cognitive Psychology, 23*(8), 922–930. https://doi.org/10.1080/20445911.2011.586780

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology, 95*(1), 57–79. https://doi.org/10.1348/000712604322779460

Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and 'theory of mind'in an agrammatic aphasic patient. *Current Biology, 10*(12), 723–726.

Vignolo, L. A. (1999). Disorders of conceptual thinking in aphasia. In G. Denes & L. Pizzamiglio (Eds.), *Handbook of clinical and experimental neuropsychology* (p. 273e288). Psychology Press.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27*(1), 92–114. https://doi.org/10.1037/0096-1523.27.1.92

Vygotsky, L. (1962). Thought and language. In *Thought and language /*. Massachusetts Institute of Technology. John Wiley.

Walker, P., & Cuthbert, L. (1998). Remembering visual feature conjunctions: Visual memory for shape-colour associations is object-based. *Visual Cognition, 5*(4), 409–455. https://doi.org/10.1080/713756794

Wallace, G. L., Peng, C. S., & Williams, D. (2017). Interfering with inner speech selectively disrupts problem solving and is linked with real-world executive functioning. *Journal of Speech, Language, and Hearing Research, 60*(12), 3456–3460. https://doi.org/10.1044/2017_JSLHR-S-16-0376

Waltz, J. A., Lau, A., Grewal, S. K., & Holyoak, K. J. (2000). The role of working memory in analogical mapping. *Memory & Cognition, 28*(7), 1205–1212. https://doi.org/10.3758/BF03211821

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20*(2), 158–177. https://doi.org/10.1037/h0074428

Weywadt, C. R. B., & Butler, K. M. (2013). The role of verbal short-term memory in task selection: How articulatory suppression influences task choice in voluntary task switching. *Psychonomic Bulletin & Review, 20*(2), 334–340. https://doi.org/10.3758/s13423-012-0349-0

Wickelgren, W. A. (1965a). Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 4*(1), 53–61. https://doi.org/10.1016/S0022-5371(65)80067-6

Wickelgren, W. A. (1965b). Acoustic similarity and intrusion errors in short-term memory. *Journal of Experimental Psychology, 70*(1), 102–108. https://doi.org/10.1037/h0022015

Wickham, L. H. V., & Swift, H. (2006). Articulatory suppression attenuates the verbal overshadowing effect: A role for verbal encoding in face identification. *Applied Cognitive Psychology, 20*(2), 157–169. https://doi.org/10.1002/acp.1176

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104*(19), 7780–7785. https://doi.org/10.1073/pnas.0701644104

Witzel, C., & Gegenfurtner, K. R. (2011). Is there a lateralized category effect for color? *Journal of Vision, 11*(12), 16. https://doi.org/10.1167/11.12.16

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *WIREs Cognitive Science, 2*(3), 253–265. https://doi.org/10.1002/wcs.104

Zeithamova, D., & Maddox, W. T. (2007). The role of visuospatial and verbal working memory in perceptual category learning. *Memory & Cognition, 35*(6), 1380–1398. https://doi.org/10.3758/BF03193609

Zettersten, M., & Lupyan, G. (2020). Finding categories through words: More nameable features improve category learning. *Cognition, 196,* 104135. https://doi.org/10.1016/j.cognition.2019.104135

Zhong, W., Li, Y., Li, P., Xu, G., & Mo, L. (2015). Short-term trained lexical categories produce preattentive categorical perception of color: Evidence from ERPs: Categories produce categorical perception. *Psychophysiology, 52*(1), 98–106. https://doi.org/10.1111/psyp.12294

Zhou, K., Mo, L., Kay, P., Kwok, V. P. Y., Ip, T. N. M., & Tan, L. H. (2010). Newly trained lexical categories produce lateralized categorical perception of color. *Proceedings of the National Academy of Sciences, 107*(22), 9974–9978. https://doi.org/10.1073/pnas.1005669107

Zourbanos, N., Hatzigeorgiadis, A., Bardas, D., & Theodorakis, Y. (2013). The effects of self-talk on dominant and nondominant arm performance on a handball task in primary physical education students. *The Sport Psychologist, 27*(2), 171–176. https://doi.org/10.1123/tsp.27.2.171

Journal : Large 13423 | Article No : 2144 | Pages : 25 | MS Code : PBR-TR-22-013 | Dispatch : 9-7-2022

Journal:	**13423**
Article:	**2144**

Author Query Form

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query	Details Required	Author's Response
AQ1	Please check "Affiliation" if captured correctly.	
AQ2	The citation "Lupyan, 2012" has been changed to "Lupyan, 2012a, b" to match the author name/date in the reference list. Please check if the change is fine in this occurrence and modify the subsequent occurrences, if necessary.	
AQ3	Experiment 1-4 has been cited in the body but no data please check and provide.	
AQ4	In jobsheet ESM reflect was NO.	