

1 Not everybody has an inner voice: Initial exploration of anendophasia and four behavioral  
2 experiments

3 Johanne S. K. Nedergaard<sup>1</sup> & Gary Lupyan<sup>2</sup>

4 <sup>1</sup> Department of Linguistics, Cognitive Science and Semiotics, Aarhus University

5 <sup>2</sup> Department of Psychology, University of Wisconsin-Madison

6 Author Note

7 All experiment data, experiment code, and analysis code are available on GitHub:  
8 <https://github.com/johannenedergaard/anendophasia>.

9 Correspondence concerning this article should be addressed to Johanne S. K.  
10 Nedergaard, Jens Chr. Skous Vej 2, 8000 Aarhus C, Denmark. E-mail: nedergaard@cc.au.dk

11

## Abstract

12 Inner speech appears to play an important role in a variety of cognitive processes but the  
13 existence of people who experience little or no inner speech in their everyday life may limit  
14 conclusions we can make about the functions of inner speech. In this exploratory study, we  
15 test two groups of participants (one reporting high levels of internal verbal representations  
16 and one reporting low levels of internal verbal representations) across four behavioral  
17 experiments (verbal working memory, rhyme judgment, same/different judgment, and task  
18 switching). These experiments were chosen to test internal verbal representations, verbal  
19 self-cuing, and the influence of inner speech on category processing. We found that  
20 participants who report low levels of verbal representations performed worse at the verbal  
21 working memory and the rhyme judgments and were slower at same/different judgments.  
22 Both groups performed equally well on the task switching experiment, however. We propose  
23 a new name for the phenomenon of experiencing little to no inner speech - *anendophasia* -  
24 and discuss the implications of our findings for the role of inner speech in cognition.

25        *Keywords:* inner speech, rhyme judgments, categorization, task switching, verbal  
26 working memory, individual differences

27        Word count: 7927

<sup>28</sup> “To think means to speak to oneself... hence to hear oneself inwardly.”

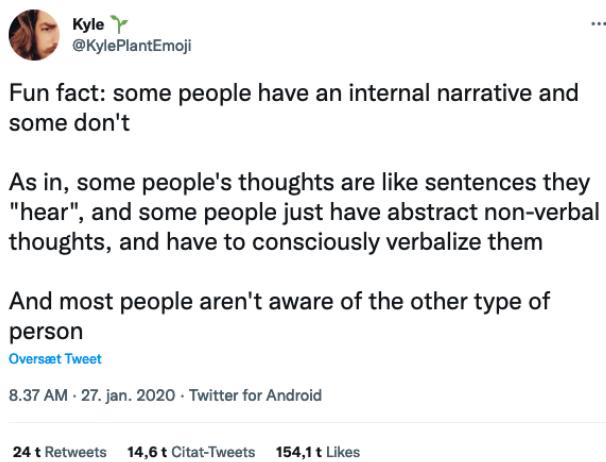
<sup>29</sup> — Immanuel Kant in *Anthropology from a pragmatic viewpoint* (1798)

<sup>30</sup> “He who is capable of thinking with the aid of language alone, has not yet  
<sup>31</sup> experienced abstract and genuine thinking.”

<sup>32</sup> — Eugen Karl Dühring as quoted in Engels (1877)

<sup>33</sup> **1 Introduction**

<sup>34</sup> Everyone has an inner voice, and most of our waking hours are filled with internal  
<sup>35</sup> monologue. These are claims that scientists of the mind casually make often (Chella &  
<sup>36</sup> Pipitone, 2020; Perrone-Bertolotti, Rapin, Lachaux, Baciu, & Loevenbruck, 2014) but the  
<sup>37</sup> picture is muddied by the recent “discovery” that some people apparently do not experience  
<sup>38</sup> an inner voice. The topic has received much attention in viral Twitter threads (e.g.,  
<sup>39</sup> @KylePlantEmoji (2020), see Figure 1) as well as in articles such as ‘What it’s like living  
<sup>40</sup> without an inner voice’ (Soloducha, 2020) and ‘People With No Internal Monologue Explain  
<sup>41</sup> What It’s Like In Their Head’ (Felton, 2020).



*Figure 1.* Viral tweet from @KylePlantEmoji about the presence or absence of inner speech.  
Screenshot from November 17th 2022.

42 Judging by these accounts, there are important differences in the extent to which  
43 people experience an inner voice. Whether these differences in experience result in differences  
44 in behavior is still an open question which we hope begin to address in the present article.

45 If people's reports are accurate and an absence or diminished awareness of inner speech  
46 is more widespread than previously thought, this has important consequences for the  
47 relationship between language/inner speech and thought. The assumption that everyone has  
48 an inner voice has served as a stepping stone for much research into the functions of inner  
49 speech – if everyone has it, it must be important. This importance ranges from claims that  
50 language constitutes (at least some types of) thought (Bermúdez, 2007; Carruthers, 2002;  
51 Clark, 1998; Frankish, 2018; Gauker, 2011; Morin, 2018) to investigations into the connection  
52 between inner speech and behavioral control (Alderson-Day & Fernyhough, 2015; Cragg &  
53 Nation, 2010; Emerson & Miyake, 2003; Morin, Duhnych, & Racy, 2018). As the two quotes  
54 at the beginning of this article suggest, philosophers and psychologists have disagreed about  
55 the role of inner speech and language in thinking for centuries. It is possible that these  
56 disagreements stem at least in part from the philosophers' and psychologists' subjective  
57 experience as this is also the case with aphantasia, the inability to engage in visual imagery  
58 (Galton, 1880). With globalization and social media has come the opportunity to discuss  
59 these intuitions of a select few thinkers with a much wider range of people whose intuitions  
60 we would do well to take seriously.

### 61 1.1 Parallels with unsymbolized thinking

62 Anecdotally, some people with no inner speech report that their thinking takes place  
63 largely in the visuospatial modality while another common description is that they 'think in  
64 concepts'. What it means to think in concepts without relying on language is contentious but  
65 the idea to some extent parallels findings from Descriptive Experience Sampling (DES)  
66 (Heavey & Hurlburt, 2008; Russell T. Hurlburt & Akhter, 2006), where participants often  
67 report what is called 'unsymbolized thinking' (around 22 % of experience prompts). In such

68 episodes, people feel that they think ‘a particular, definite thought without the awareness of  
69 that thought’s being conveyed words, images, or any other symbols’ (Heavey & Hurlburt,  
70 2008, p. 802). Unsymbolized thinking is a slippery phenomenon mostly characterized with  
71 negative definitions. For example, Russell T. Hurlburt and Akhter (2008) say that it is  
72 experienced as being ‘a thinking, not a feeling, not an intention, not an intimation, not a  
73 kinesthetic event, not a bodily event’ (p. 1366). The existence of conscious, unsymbolized  
74 thinking poses a challenge to views that hold that language is necessary for conscious  
75 thinking, as does the putative existence of people who do not experience inner speech at all.  
76 Indeed, proponents of such views tend to deny the existence of unsymbolized thought,  
77 claiming that reports may be a result of confabulation (Carruthers, 2009). Either people  
78 who report such thoughts are using words or images without being aware of it, or they are  
79 not really thinking (e.g. confusing looking at something with considering something). One  
80 way of testing the confabulation idea is to test whether the presence or absence of verbal  
81 thought has behavioral consequences which is what we did in the series of experiments  
82 discussed in the present paper.

### 83 1.2 Parallels with aphantasia

84 That there are differences in subjective reports of inner experience is not a new finding,  
85 nor is the idea that such differences may result in subtle behavioral changes. In recent years,  
86 a very similar phenomenon to internal verbal experience has gained much attention, namely  
87 the presence or absence of visual imagery. In a 2010 article, Zeman and colleagues termed  
88 the inability to engage in visual imagery ‘aphantasia’ and reported that two thirds of the  
89 participants had difficulties with autobiographical memory as a result of their aphantasia  
90 (Zeman et al., 2010). Generally, participants with aphantasia report weak or non-existing  
91 ability to visualize ‘in the mind’s eye’ (Dawes, Keogh, Andrillon, & Pearson, 2020; Keogh &  
92 Pearson, 2018) and may display poorer visual working memory performance than control  
93 participants (Jacobs, Schwarzkopf, & Silvanto, 2018) although this is not always the case

94 (Keogh, Wicken, & Pearson, 2021). The conflicting findings about consequences of  
95 aphantasia in terms of working memory abilities have prompted a discussion of whether  
96 aphantasia represents a metacognitive deficit rather than difficulties with mental visual  
97 imagery. However, recent findings suggest that a more likely explanation is that people with  
98 aphantasia simply use different strategies to solve tasks that would normally require visual  
99 imagery. For example, Keogh, Wicken, and Pearson (2021) found that participants with  
100 aphantasia performed at the same level as control participants on visual working memory  
101 tasks. There were, however, marked differences in the reported strategies used by  
102 participants with aphantasia who reported rehearsing patterns verbally or ‘using ideas and  
103 semantics’ to remember the test items. Additionally, performance levels on a number  
104 working memory task and a visual working memory task were correlated for participants  
105 with aphantasia but not for control participants. This suggests that control participants used  
106 different strategies for the two types of tasks (one is traditionally thought to occupy verbal  
107 resources while the other is thought to utilize visual working memory resources) while  
108 participants with aphantasia may have used similar strategies for the two different tasks.  
109 The finding that differences in strategies are likely to mask differences in visualizing ability is  
110 important for research in inner speech as well. We might see comparable performance levels  
111 due to compensatory strategies that would then mask differences in mental verbalizing  
112 abilities.

113 There are, however, also important differences between aphantasia and lack of inner  
114 speech as inner speech cannot be reduced to auditory imagery. First, inner speech involves  
115 both auditory and articulatory-motor imagery (Geva, 2018; Perrone-Bertolotti et al., 2014).  
116 Second, while internal language addressed to oneself is often experienced as having  
117 phonological features - one of the reasons people often perceive it as speech  
118 (Langland-Hassan, 2018) - it does not necessarily. For example, inner speaking and inner  
119 signing have something in common as ways of producing internal language addressed to  
120 oneself that has nothing to do with auditory imagery (McGuire et al., 1997).

<sup>121</sup> **1.3 How widespread might a lack of inner speech be?**

<sup>122</sup> It is difficult to assess the prevalence of inner speech in the general population but  
<sup>123</sup> there are nevertheless several sources for making educated guesses. People with aphantasia  
<sup>124</sup> often report difficulties with imagery in other modalities as well (Dawes et al., 2020) and so  
<sup>125</sup> we might expect a similar incidence rate for lack of inner speech. If we take this as a starting  
<sup>126</sup> point, we could see little to no inner speech in around 2-4 % of the population (Dance, Ipser,  
<sup>127</sup> & Simner, 2022; Faw, 2009). Alternatively, we could look at what proportion of children  
<sup>128</sup> engage in private speech (93 % according to Winsler, De Leon, Wallace, Carlton, and  
<sup>129</sup> Willson-Quayle (2003)) or how many adult participants report never engaging in self-talk in  
<sup>130</sup> large-scale questionnaire studies (e.g. less than 10 % in Nedergaard, Christensen, and  
<sup>131</sup> Wallentin (2021)). Regardless, people with no inner speech appear to be in a small minority,  
<sup>132</sup> and the phenomenon is likely to be continuous rather than dichotomous.

<sup>133</sup> **1.4 Which behavioral consequences of less inner speech would we expect?**

<sup>134</sup> Say that we take people's claims seriously when they say that they experience no inner  
<sup>135</sup> speech. Given the myriad of theories on the role that language plays in cognition, we would  
<sup>136</sup> expect them to also think differently. For example, we might believe that inner speech is  
<sup>137</sup> recruited for (or actually constitutes) self-reflection and metacognition (Morin, 2018; Clark,  
<sup>138</sup> 1998), and so people with less inner speech should perform less well on metacognitive tasks.  
<sup>139</sup> If we find both that they are right about not having inner speech and that they perform  
<sup>140</sup> equally well, then we might conclude that inner speech is probably not crucial for  
<sup>141</sup> metacognition. This line of reasoning hinges on finding a good way to test whether people  
<sup>142</sup> are right when they claim that they experience no inner speech and finding tests of cognitive  
<sup>143</sup> abilities that either could only be solved with inner speech or would allow us to differentiate  
<sup>144</sup> between verbal and non-verbal solutions/strategies (despite them possibly resulting in  
<sup>145</sup> comparable performance).

<sup>146</sup> **1.5 The present study**

<sup>147</sup> To test whether participants are accurate about their internal verbal representations,  
<sup>148</sup> we use a rhyme judgment task (Geva, Bennett, Warburton, & Patterson, 2011;  
<sup>149</sup> Langland-Hassan, Faries, Richardson, & Dietz, 2015) where participants see two images and  
<sup>150</sup> have to judge whether the associated words rhyme or not. Presumably, this would require  
<sup>151</sup> them to internally verbalize. Importantly, we need to include both orthographic rhymes  
<sup>152</sup> (such as ‘boat’ and ‘moat’) and non-orthographic rhymes (such as ‘sleigh’ and ‘hay’) as  
<sup>153</sup> participants could otherwise make rhyme judgments by visualizing the orthographic  
<sup>154</sup> representations of the words. In functional neuroimaging studies, rhyme judgment tasks are  
<sup>155</sup> associated with activation in language production areas such as the left inferior frontal gyrus  
<sup>156</sup> and the inferior parietal lobe (Hoeft et al., 2007; Lurito, Kareken, Lowe, Chen, & Mathews,  
<sup>157</sup> 2000; Owen, Borowsky, & Sarty, 2004; Paulesu, Frith, & Frackowiak, 1993; Poldrack et al.,  
<sup>158</sup> 2001; Pugh et al., 1996). Taking inspiration from aphantasia research on visual working  
<sup>159</sup> memory, it also seems reasonable to focus on verbal working memory in the case of people  
<sup>160</sup> with little to no inner speech. In particular, we might expect difficulties with verbal working  
<sup>161</sup> memory tasks requiring a high degree of phonological precision (Jacobs et al., 2018). For  
<sup>162</sup> this, we use verbal working memory tests with orthographically similar sets, phonologically  
<sup>163</sup> similar sets, and control sets.

<sup>164</sup> There is robust evidence that inner speech is generally recruited for behavioral control  
<sup>165</sup> in task switching paradigms where participants have to switch between different task rules  
<sup>166</sup> (Baddeley, Chincotta, & Adlam, 2001; Emerson & Miyake, 2003; Miyake, Emerson, Padilla,  
<sup>167</sup> & Ahn, 2004). Here, participants for example have to switch between adding and subtracting  
<sup>168</sup> numbers while also performing simultaneous tasks designed to interfere (see Nedergaard,  
<sup>169</sup> Wallentin, and Lupyan (2022) for a systematic review of the verbal interference literature).  
<sup>170</sup> People who do not habitually use inner speech might resort to other means of self-cueing in  
<sup>171</sup> such cases. Given the robust literature on inner speech being involved in task switching, we

172 also test participants' ability to switch between simple addition and subtraction problems.  
173 This helps shed light on the consequences of (lack of) inner speech for behavioral control.

174 Aside from behavioral control, there is also empirical evidence that language influences  
175 thought on a more structural level through imposing categories on for example the color  
176 spectrum (Gilbert, Regier, Kay, & Ivry, 2006, 2008; Winawer et al., 2007). Given that even  
177 people who report no inner speech grow up using language and learning categories through  
178 language, it is perhaps more of an open question whether we would expect them to be  
179 influenced by categories to the same extent as people with a lot of inner speech. Related to  
180 this idea that language structures the way we think about categories, we also test  
181 participants' ability to detect simple visual differences as well as categorical differences  
182 between line drawings of cats and dogs.

183 **2 Methods**

184 **2.1 Participants**

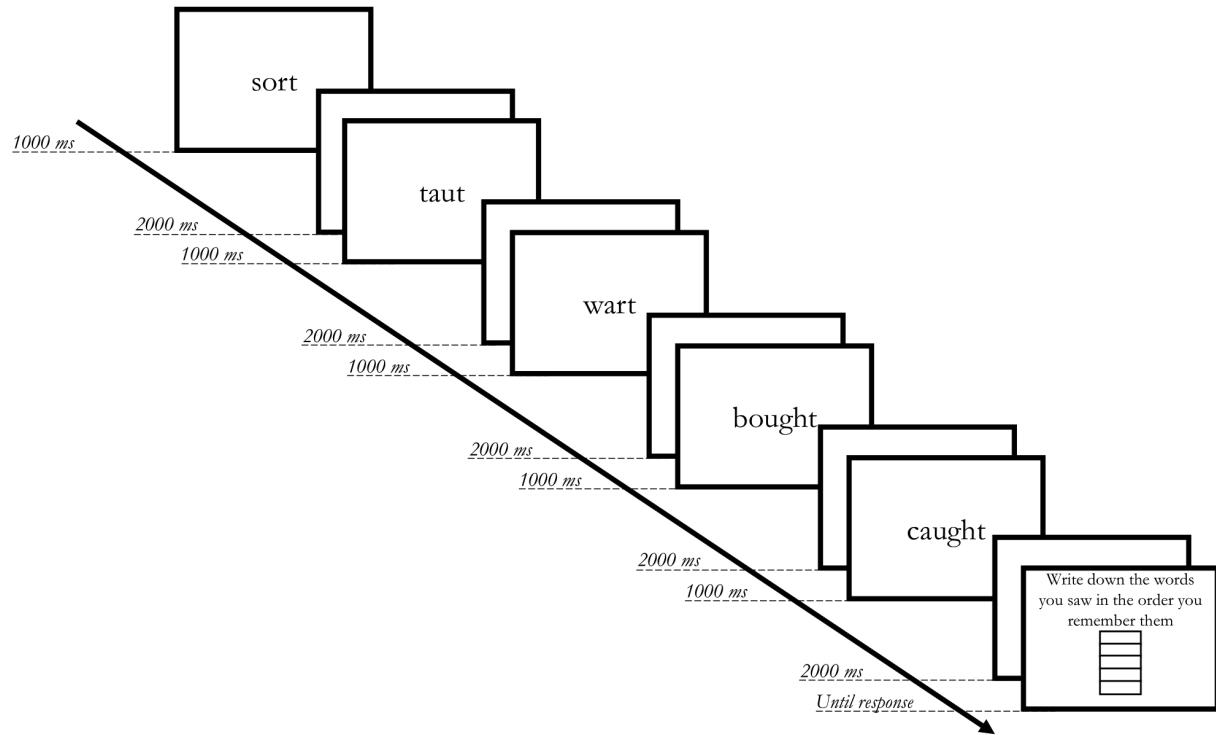
185 We contacted participants who had previously completed the Internal  
186 Representations Questionnaire (Roebuck & Lupyan, 2020) and recruited ones  
187 with the lowest and highest scores on verbal representation. This study  
188 received ethical approval from XXXX. ((Gary?) how exactly was this done?) We will  
189 use the shorthand 'high-verbal' and 'low-verbal' to refer to these two groups although it  
190 should be noted that these just refer to their inclination or propensity to engage in verbal  
191 mental imagery and not to their general linguistic abilities. All four experiments were  
192 conducted using custom-written software with the JavaScript package jsPsych version 6 (De  
193 Leeuw, 2015).

194 **2.2 Method: Verbal working memory**

195 **2.2.1 Materials.** We used the same three word sets as Baddeley (1966): One set  
196 contained words that were phonologically similar but not orthographically similar ('bought',

197 ‘sort’, ‘taut’, ‘caught’, and ‘wart’), one set contained words that were orthographically similar  
198 but not phonologically similar (‘rough’, ‘cough’, ‘through’, ‘dough’, ‘bough’), and one set was  
199 a control set (‘plea’, ‘friend’, ‘sleigh’, ‘row’, ‘board’).

200       **2.2.2 Procedure.** Participants received the following instructions: ‘During each  
201 trial, you will see five words presented in sequence 1 second each. You will have to remember  
202 the words and the order they were presented in as you will be asked to reproduce them in  
203 the right order afterwards.’ There was a blank screen for 2000 ms between each word  
204 presentation. First, participants performed two practice trials with full feedback  
205 (correct/incorrect and the stimulus words – drawn from a different set than the ones used in  
206 the real experiment – shown in order). Then, participants performed 24 trials with eight  
207 trials from each of the three word sets. The order of both set type and words within a trial  
208 were randomized. There was no limit to how long participants could spend on reproducing  
209 the words on a given trial. See Figure 2.



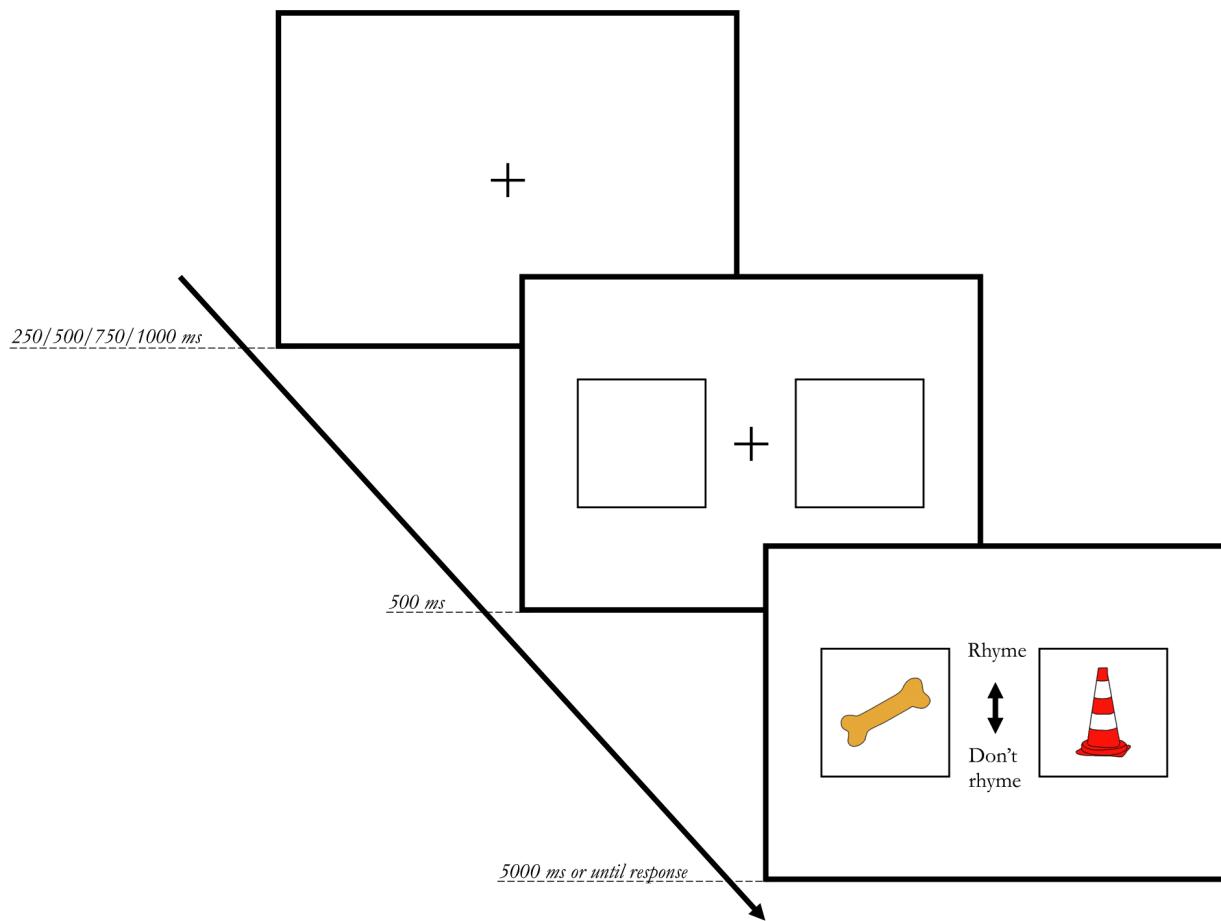
*Figure 2.* A sketch of the procedure in the verbal working memory experiment. In this example, the words are drawn from the phonological similarity set

<sup>210</sup> **2.3 Method: Rhyme judgments**

<sup>211</sup> **2.3.1 Materials.** We constructed a set of rhyme pairs with 20 orthographical pairs  
<sup>212</sup> (e.g. ‘sock’ and ‘clock’) and 20 non-orthographical pairs (e.g. ‘drawer’ and ‘door’). See  
<sup>213</sup> Supplementary Materials for the full set of images, associated words, and name agreement  
<sup>214</sup> scores. The images were selected from the MultiPic database (Duñabeitia et al., 2018) and  
<sup>215</sup> from Snodgrass and Vanderwart (1980).

<sup>216</sup> **2.3.2 Procedure.** Participants received the following instructions: ‘You will see  
<sup>217</sup> two images at a time and have to judge whether the names of the items rhyme or not. For  
<sup>218</sup> example, if you see a picture of a LAMP and a picture of a CAMP, you should respond that  
<sup>219</sup> they rhyme (press UP arrow). If you see a picture of a BEAR and a picture of a CUP, you  
<sup>220</sup> should respond that they do not rhyme (press DOWN arrow). All the words are short (one

221 syllable). Please make the judgments as quickly and accurately as possible.' Participants first  
 222 performed four practice trials with correct/incorrect feedback – they did not receive feedback  
 223 for the remaining trials. Between each rhyme judgment trial, the screen showed a central  
 224 fixation cross for either 250, 500, 750, or 1000 ms. It then showed two square black frames  
 225 for 500 ms to control spatial attention – the two images then appeared simultaneously in the  
 226 two squares. Participants had 5000 ms to respond to each trial and performed a total of 60  
 227 rhyme judgments in randomized order (20 orthographical rhymes, 20 non-orthographic  
 228 rhymes, and 20 no-rhyme control trials). See Figure 3.



*Figure 3.* A sketch of a rhyme judgment trial. The stimuli here exemplify an orthographic rhyme – ‘bone’ and ‘cone’ – and the correct answer would therefore be ‘Rhyme’.

229 **2.4 Method: Task switching**

230       **2.4.1 Materials.** For each of the five experimental conditions, we used 30

231 randomly selected integers between 13 and 96 as prompts to make sure that all correct

232 results were two-digit positive numbers.

233       **2.4.2 Procedure.** There were five conditions in this experiment: (1) blocked

234 addition (2) blocked subtraction, (3) alternating between addition and subtraction with

235 operation marked by color cue (red/blue); (4) alternation marked with a symbol cue (+/-);

236 (5) alternation without external cue requiring participants to remember which operation they

237 just did.. Participants started with either the blocked subtraction or the addition condition

238 (counterbalanced) and then proceeded to the switching conditions (counterbalanced). For

239 each condition, participants first solved 10 problems with correct/incorrect feedback

240 (including feedback specific to whether the arithmetic or the operation or both were

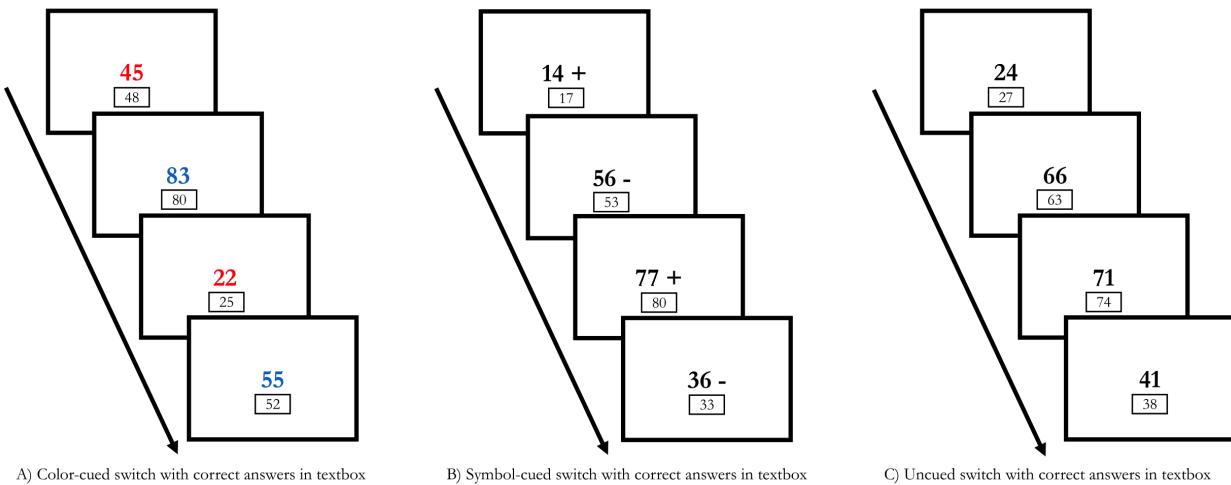
241 incorrect) and then 30 problems without feedback. On a given trial, participants saw a

242 prompt number and had to either add or subtract 3 and type their answer into a text box.

243 In the switching conditions, a response counted as correct if it was the correct arithmetic and

244 if the operation was switched from the previous trial (from addition to subtraction or vice

245 versa). See Figure 4.



*Figure 4.* A sketch of the three switched conditions in the task switching experiment. Figure A shows four color-cued switch trials with correct answers, Figure B shows four symbol-cued switch trials with correct answers, and Figure C shows four uncued switch trials with correct answers.

<sup>246</sup> **2.5 Method: Same/different judgments**

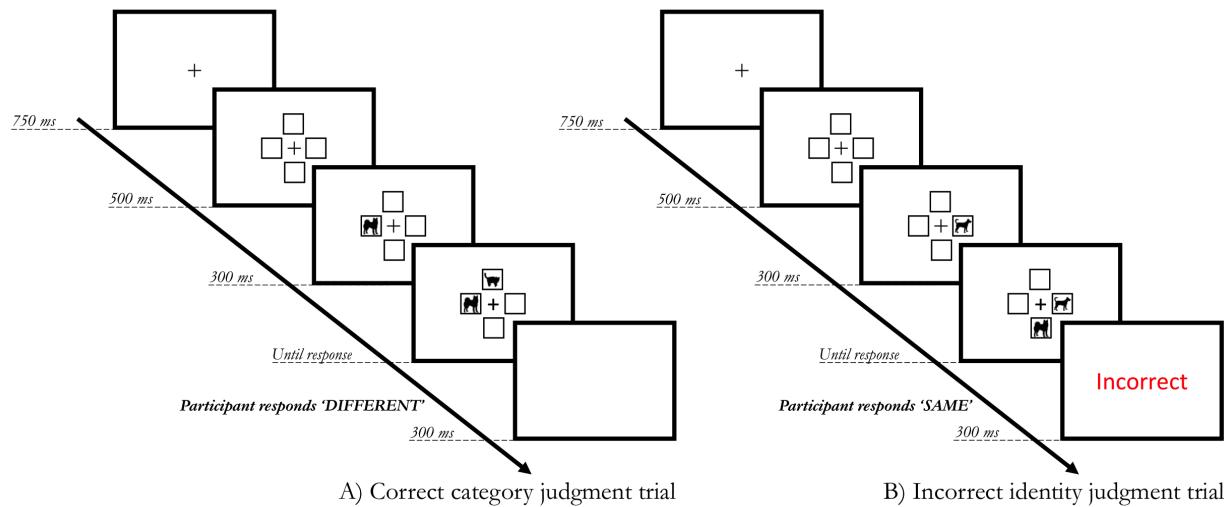
<sup>247</sup> **2.5.1 Materials.** This experiment used three different black silhouettes of cats  
<sup>248</sup> and three different black silhouettes of dogs (see Figure 5).



*Figure 5.* The black silhouettes of cats and dogs used in the same/different judgment experiment.

<sup>249</sup> **2.5.2 Procedure.** There were two conditions in the experiment: a category  
<sup>250</sup> judgment condition and an identity judgment condition. In the category judgment condition,  
<sup>251</sup> participants were instructed to press the UP arrow key if the two animals belonged to the  
<sup>252</sup> same category (either cat or dog) and the DOWN arrow key if they did not. In the identity

judgment condition, participants were instructed to press the UP arrow key if the two animals were completely identical (e.g., same silhouette of same dog) and the DOWN arrow key if they were not. See Figure 6. On each trial, participants first saw a fixation cross for 750 ms, then four empty square frames around the fixation cross for 500 ms to prompt participants' spatial attention. The silhouette images appeared one at a time with a 300 ms delay between them in two out of four random positions around a fixation cross in the center of the screen. After the keyboard response, the screen was blank for 300 ms. Participants received feedback throughout but only for incorrect trials. They completed 100 trials in the category judgment condition and 100 trials in the identity judgment condition (half 'same' and half 'different').



*Figure 6.* A sketch of the two conditions of the category judgment experiment. On Figure A, we see a correct category judgment trial where the participant responds that the cat and dog silhouettes represent different animals. On Figure B, we see an incorrect identity judgment trial where the participant responds that the two dogs are identical.

## 2.6 Method: Questionnaire

After completing the four experiments, participants answered the following custom questions. They also completed the Varieties of Inner Speech Questionnaire (VISQ)

<sup>266</sup> (Alderson-Day, Mitrenga, Wilkinson, McCarthy-Jones, & Fernyhough, 2018).

Question	Options
If you have to ask a question in front of an audience, which of these best describes what you typically do?	I rehearse in my mind the exact phrasing of what I am going to ask (5) I rehearse in my mind some of what I am going to ask before asking it (4) I think of a question I want to ask and just ask it (3) Other (2)
How often do you experience trouble focusing on a face-to-face conversation you are having because of a conflicting conversation happening in your mind at the same time?	I'm never in a position to ask questions in front of an audience (1) Never (1) Rarely (2) Sometimes (3) Often (4) Always (5)
How often do you have songs stuck in your head?	Multiple times a day (5) A few times a week (4) A few times a month (3) A few times a year (2) Never (1)
If you had to recall a short conversation about a specific topic that you had yesterday with a friend, how easily can you recall the exact words your friend said?	I can easily recall it. If I wrote it down and matched to a recording of the conversation, there'd be an almost perfect match (5) I remember the topic and remember much of what was said. If I matched it to a recording of the conversation, a lot would match up. (4) I remember the topic, but remember only a few of the specific words/sentences. (3) I remember the topic, but can't remember any of the specifics. (2) Other (1)

*(continued)*

Question	Options
If you had to recall a short conversation about a specific topic that you had yesterday with a friend, how easily can you recall the exact words you said?	I can easily recall it. If I wrote it down and matched to a recording of the conversation, there'd be an almost perfect match (5)  I remember the topic and remember much of what was said. If I matched it to a recording of the conversation, a lot would match up. (4)  I remember the topic, but remember only a few of the specific words/sentences. (3)  I remember the topic, but can't remember any of the specifics. (2)  Other (1)  It's just like I'm hearing the conversation again. (4)
When you recall a conversation like the one you were thinking about for the last 2 questions, do you hear the words in your mind?	I hear a condensed version (e.g. only some words). (3)  I hear something but I can't describe it. (2)  I can't hear it, but I can still recall it. Please briefly say something about how you are recalling it. (1)
Can you "sing along" to music without singing out loud?	Yes - definitely (4)  Yes - somewhat (3)  No - but I can imagine how others can do it (2)  No - I can't imagine how anyone could do this (1)
If you can "sing along" to music without singing out loud, to what extent does this feel like regular thinking?	Not at all (1)  Mostly different from regular thinking (2)  Neutral (3)  Mostly similar to regular thinking (4)  Exactly like regular thinking (5)  I can't sing along without singing out loud (6)
If you imagine someone else speaking, how do you experience their voice?	I hear what they say in their voice. (4)  I hear what they say but in my own voice. (3)  I hear the words but I can't tell whose voice it is. (2)  I don't "hear" anything, I imagine it by... (please specify) (1)
Many people feel that a lot of their thinking, planning, and decision-making takes place in the form of a conversation with themselves. They describe that when they think, they hear words in their mind. Other people don't have this experience and instead say that they "think in ideas". Is your experience more like the first or the second?	More like a conversation (2)  More like "thinking in ideas". Can you elaborate or give an example of what this means to you? (1)

*(continued)*

Question	Options
To what extent do you agree with this statement: 'It is generally difficult and takes effort to express in words how I think and feel'.	Strongly agree (1) Agree (2) Neither agree nor disagree (3) Disagree (4) Strongly disagree (5)
Do you think it is stressful and annoying to have an inner monologue?	Yes, very (3) Maybe a little (2) No, I don't think so (1)
In books and movies, we often see characters talking to themselves at length. How much do you think this reflects real life?	It's just for the viewer/reader's benefit (1) It might be like real life but mostly for the viewer's/reader's benefit (2) It's exactly like real life (3)
Have you been diagnosed with dyslexia or another reading disorder?	Yes, officially diagnosed (1) Yes, self-diagnosed (2) No, never (3)
Do you ever revise past conversations in your mind (i.e. think of a better comeback, a way of phrasing what you wanted to say)?	Never (1) Rarely (2) Sometimes (3) Often (4) Very often (5)
Do you ever rehearse a conversation before you have it in real life where you simulate what you will say and how the other person will respond?	Never (1) Rarely (2) Sometimes (3) Often (4) Very often (5)
Imagine you are lying in bed with your eyes closed trying to fall asleep. Is your inner experience then...	Primarily verbal (you "hear" or "speak" words and sentences in your mind) (1) Primarily visual (you "see" situations, objects, people etc. in your mind) (2) Primarily about sensory awareness (what you are hearing, smelling, and feeling in the moment) (3) Primarily emotional (4) An even mix of verbal, visual, sensory, and emotional (5) My inner experience in that situation does not have a specific "format" (6)

(continued)

Question	Options
To what extent do you agree with this statement: "I do not know why I do some of the things that I do."	Strongly disagree (1) Disagree (2) Neither agree nor disagree (3) Agree (4) Strongly agree (5)
To what extent do you agree with this statement: "I am a firm believer in thinking things through."	Strongly disagree (1) Disagree (2) Neither agree nor disagree (3) Agree (4) Strongly agree (5)
To what extent do you agree with this statement: "I like to act on a whim."	Strongly disagree (1) Disagree (2) Neither agree nor disagree (3) Agree (4) Strongly agree (5)
For each scale, please indicate what percent of people you know you think have each of these three experiences:	No one (0%) to Everyone (100%)
- Experience their thoughts in the form of a conversation with themselves - Can see vivid images in their mind's eye - Hear words in their mind's ear when they silently read	

## 267 2.7 Data analysis

268 We used R (Version 4.1.3; R Core Team, 2022) and the R-packages *corrplot*2021 (Wei  
 269 & Simko, 2021), *cowplot* (Version 1.1.1; Wilke, 2020), *data.table* (Version 1.14.0; Dowle &  
 270 Srinivasan, 2021), *dplyr* (Version 1.0.7; Wickham, François, Henry, & Müller, 2021), *forcats*  
 271 (Version 0.5.1; Wickham, 2021a), *Formula* (Version 1.2.4; Zeileis & Croissant, 2010), *ggforce*  
 272 (Version 0.3.3; Pedersen, 2021), *ggplot2* (Version 3.3.5; Wickham, 2016), *ggpubr* (Version  
 273 0.4.0; Kassambara, 2020), *Hmisc* (Version 4.5.0; Harrell Jr, Charles Dupont, & others.,  
 274 2021), *kableExtra* (Version 1.3.4; Zhu, 2021), *lattice* (Version 0.20.45; Sarkar, 2008), *lme4*  
 275 (Version 1.1.27.1; Bates, Mächler, Bolker, & Walker, 2015), *lmerTest* (Version 3.1.3;  
 276 Kuznetsova, Brockhoff, & Christensen, 2017), *Matrix* (Version 1.4.0; Bates & Maechler,

277 2021), *optimx* (Nash, 2014; Version 2021.10.12; Nash & Varadhan, 2011), *papaja* (Version  
278 0.1.1; Aust & Barth, 2022), *purrr* (Version 0.3.4; Henry & Wickham, 2020), *readr* (Version  
279 2.1.1; Wickham, Hester, & Bryan, 2021), *stringr* (Version 1.4.0; Wickham, 2019), *survival*  
280 (Version 3.2.13; Terry M. Therneau & Patricia M. Grambsch, 2000), *tibble* (Version 3.1.6;  
281 Müller & Wickham, 2021), *tidyverse* (Version 1.3.1; Wickham et al., 2019), *tinylabels* (Version 0.2.3; Barth, 2022), *trackdown* (Version 1.1.1;  
282 Wickham et al., 2019) for all our analyses. In all linear mixed models, we modelled participant as  
283 random intercepts as well as random intercepts for all within-subjects factors (e.g. condition  
284 with repeated measures). For the statistical models predicting reaction time, we  
285 log-transformed the dependent variable to prevent issues with non-normal distributions. For  
286 the statistical models predicting accuracy, we constructed binomial generalized linear mixed  
287 models. All the plots visualize categorical differences between the two groups while all the  
288 statistical models use verbal score (average score on the verbal representation items on the  
289 Internal Representations Questionnaire) as a continuous predictor.

290       **2.7.1 Exclusion criteria.** We excluded 10 participants for responding randomly,  
291 missing at least one out of the four experiments, or otherwise not complying with task  
292 instructions. Our final sample included 47 participants with high verbal representation  
293 scores on the IRQ and 46 participants with low-verbal representation scores.

294

### 3 Results

295       **3.1 Verbal working memory**

296       **3.1.1 Descriptive statistics by group: Verbal working memory.**

297 High-verbal participants recalled more words correctly. This advantage was evident both  
298 when we scored only correctly ordered responses as correct as well as when we scored  
299 correctly recalled items regardless of their position (see Table 2 and Figure 7).

**Table 2**

*Descriptive statistics by group in the verbal working memory experiment.*

Group	Word set	Score (item	95% CI score	Score (position	95% CI score
		and position)	(item and position)	indifferent	(position indifferent)
High-verbal	Control set	4.19	0.13	4.51	0.08
High-verbal	Orthographical similarity set	3.72	0.14	4.18	0.10
	Phonological similarity set	3.43	0.16	4.11	0.10
Low-verbal	Control set	3.69	0.15	4.17	0.11
Low-verbal	Orthographical similarity set	3.52	0.15	4.10	0.11
	Phonological similarity set	3.02	0.15	3.81	0.11

### 3.1.2 Statistical models: Verbal working memory.

We conducted two linear mixed models of original word set (phonologically similar, orthographically similar, and control set) and verbal score predicting either memory performance with both correct word and correct position or memory performance with correct word regardless of position. Both models included random intercepts by participant for the three word sets. For memory performance requiring both accurate word and position, both the set with phonologically similar words and the set with orthographically similar words were more difficult than the control set (phonological set:  $\beta = -0.72$ ; SE = 0.08;  $t = -8.84$ ;  $p < .001$ ; orthographical set:  $\beta = -0.33$ ; SE = 0.08;  $t = -3.98$ ;  $p < .001$ ). A higher verbal score was associated with increased memory performance ( $\beta = 0.27$ ; SE = 0.10;  $t = 2.60$ ;  $p = 0.01$ ). There were no interaction effects (all  $p > 0.19$ ). The same pattern was found when the correct word in any position counted as correct: The phonologically similar set was more difficult than the control set ( $\beta = -0.37$ ; SE = 0.07;  $t = -5.51$ ;  $p < .001$ ) and so was the orthographically similar set ( $\beta = -0.21$ ; SE = 0.06;  $t = -3.35$ ;  $p = 0.00$ ). A higher verbal score was associated with increased memory performance ( $\beta = 0.19$ ; SE = 0.08;  $t = 2.57$ ;  $p = 0.01$ ). There were

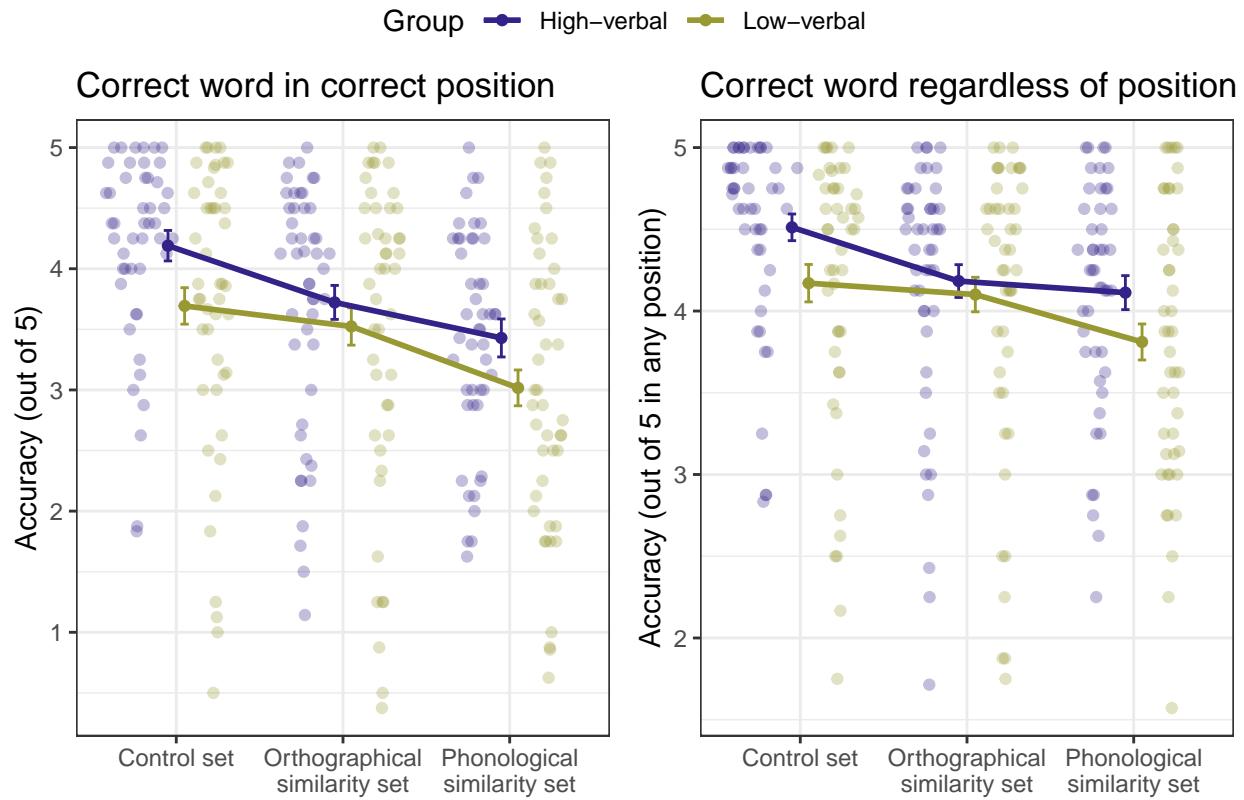
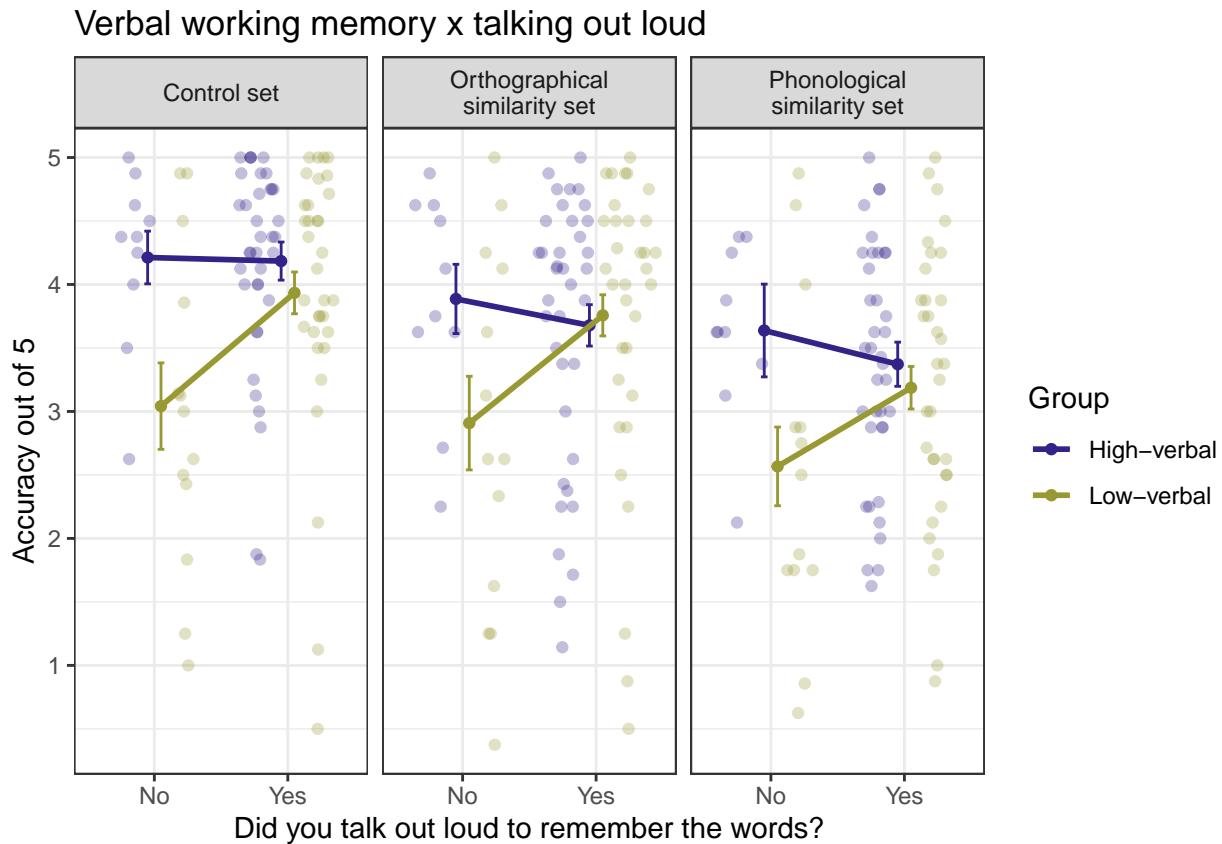


Figure 7. Score on the verbal working memory task by word set. Error bars indicate 95 % confidence intervals.

<sup>317</sup> no interaction effects (all  $p > 0.10$ ).

<sup>318</sup> **3.1.3 Strategies: Verbal working memory.** We were interested in whether <sup>319</sup> participants said the words out loud to help them remember them. We asked about this at <sup>320</sup> the end of the experiment. A chi-squared test showed that there was no significant difference <sup>321</sup> between how many high-verbal participants (10 out of 47) and how many low-verbal <sup>322</sup> participants (13 out of 46) reported that they had said the words out loud ( $\chi^2(1) = 0.29$ ,  $p$  <sup>323</sup> = 0.59). Nevertheless, the effect of doing so was interestingly different for the two groups as <sup>324</sup> can be seen in Figure 8.

<sup>325</sup> The difference between the two groups' memory performance disappears when they <sup>326</sup> report that they said the words out loud to help them remember. Doing so helps low-verbal <sup>327</sup> participants but makes no difference for high-verbal participants. Participants gave some



*Figure 8.* Verbal working memory performance by whether participants reported talking out loud to help them remember or not. Error bars indicate 95 % confidence intervals.

328 interesting alternative strategies in response to the free answer question about strategies:

329        *High-verbal group*

- 330     • Remembering the order of the first letters once the words were familiar (e.g. c, b, t, r, d for ‘cough’, ‘bough’, ‘through’, ‘rough’, ‘dough’). One participant reported this.
- 331     • Finding a cadence/melody and using this to repeat the words.
- 332     • Chunking.
- 333     • Hand and body gestures.
- 334     • Creating a story or a sentence with the words in order (both visual and verbal). This one was the most common strategy.

337       *Low-verbal group*

- 338       • Remembering the order of the first letters once the words were familiar (e.g. c, b, t, r,  
339           d for ‘cough’, ‘bough’, ‘through’, ‘rough’, ‘dough’). This strategy was much more  
340           common for the low-verbal group than for the high-verbal group.
- 341       • Form a story or a narrative. This was a less common strategy than remembering the  
342           first letters.

343       **3.2 Rhyme judgments**

344       We excluded five rhyming pairs as they had below-chance performance on average for at  
345           least one group. These pairs were bin/chin, cab/crab, rake/cake, wave/cave, and park/shark.

346       **3.2.1 Descriptive statistics by group: Rhyme judgments.** Here is a table of  
347           accuracy and reaction time for the two groups (high- and low-verbal) across types of  
348           rhyming trials.

349       As can be seen in Table 3, high-verbal participants were generally both faster and more  
350           accurate than low-verbal participants on all three types of trials. See also Figure 9.

351       **3.2.2 Statistical models: Rhyme judgments.** Both the model predicting  
352           reaction time and the model predicting accuracy included random intercepts by participant  
353           for the within-subjects variable rhyme type. All predictors were centered. A model of verbal  
354           score, rhyme type, and name agreement for the first image predicting log-transformed  
355           reaction time showed no main effect of verbal score ( $\beta = -0.01$ ; SE = 0.02;  $t = -0.63$ ;  $p =$   
356           0.53), but it did find a significant effect of no-rhyme trials being slower than other trials ( $\beta =$   
357           -0.04; SE = 0.01;  $t = -5.90$ ;  $p < .001$ ) and a significant effect of higher name agreement being  
358           associated with faster reaction times ( $\beta = -0.04$ ; SE = 0.00;  $t = -8.44$ ;  $p < .001$ ). There were  
359           no significant interactions between rhyme type and verbal score ( $p = 0.56$ ). Another model  
360           of verbal score, rhyme type, and name agreement for the first image predicting accuracy  
361           showed that no-rhyme trials were easier than other trials ( $\beta = -0.95$ ; SE = 0.09;  $z = -11.03$ ;  $p$

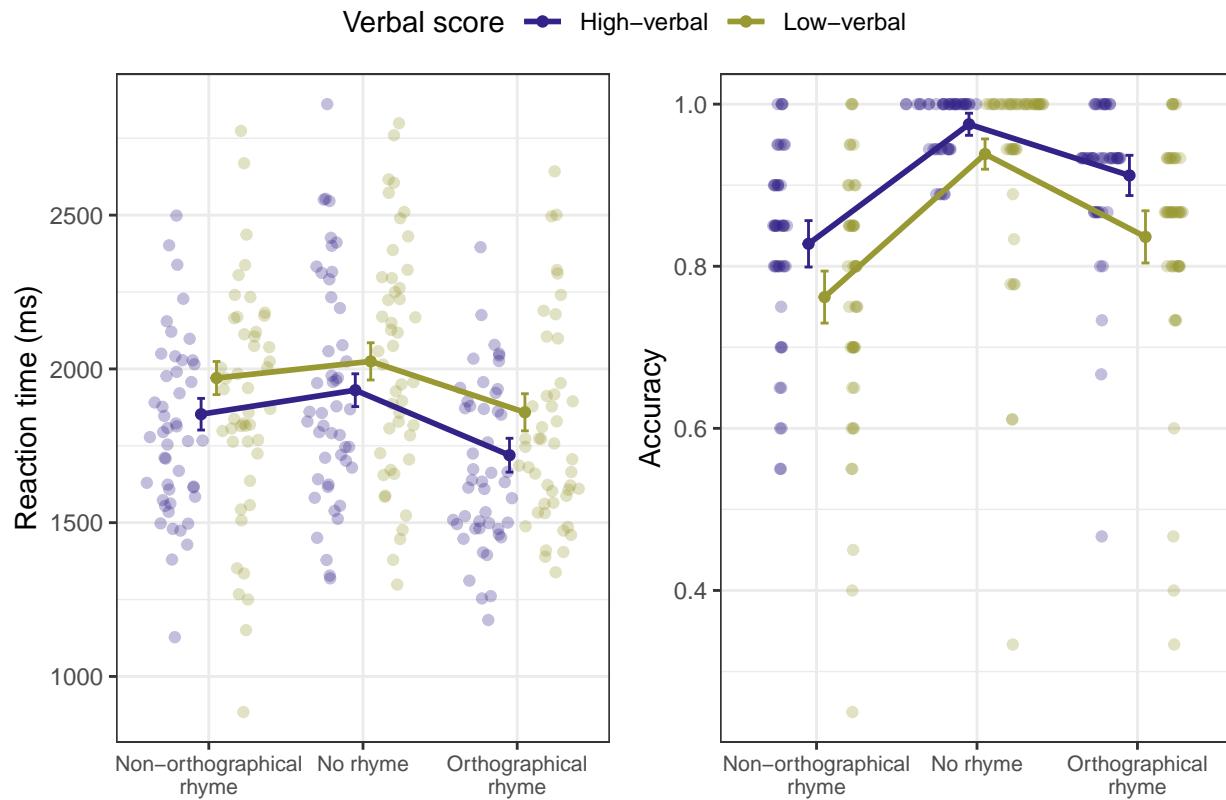
**Table 3**

*Descriptive statistics on rhyming accuracy and reaction time by group and by rhyme type.*

Group	Type of rhyme trial	Reaction time (ms)	95% CI (reaction time)	Accuracy	95% CI (accuracy)
High-verbal	Non-orthographical rhyme	1852.66	51.47	82.77	2.86
High-verbal	No rhyme	1930.79	53.26	97.52	1.36
High-verbal	Orthographical rhyme	1719.41	54.99	91.21	2.48
Low-verbal	Non-orthographical rhyme	1970.28	53.85	76.20	3.21
Low-verbal	No rhyme	2024.48	60.47	93.84	1.87
Low-verbal	Orthographical rhyme	1858.94	60.38	83.62	3.22

<sup>362</sup> < .001) and that a higher verbal score was associated with a higher likelihood of responding  
<sup>363</sup> accurately ( $\beta = 0.30$ ; SE = 0.10;  $z = 3.02$ ;  $p = 0.00$ ). It also showed that trials with images  
<sup>364</sup> with higher name agreement were significantly easier ( $\beta = 0.13$ ; SE = 0.04;  $z = 3.29$ ;  $p =$   
<sup>365</sup> 0.00). There was no significant interaction between rhyme type and verbal score ( $p = 0.40$ ).

<sup>366</sup> **3.2.3 Strategies: Rhyme judgments.** We were interested in whether  
<sup>367</sup> participants said the words out loud to make the rhyme judgments and so we included this  
<sup>368</sup> as a question at the end of the rhyming experiment. A chi-squared test showed that there  
<sup>369</sup> was no significant difference between how many high-verbal participants (23 out of 47) and  
<sup>370</sup> how many low-verbal participants (21 out of 46) reported that they had said the words out



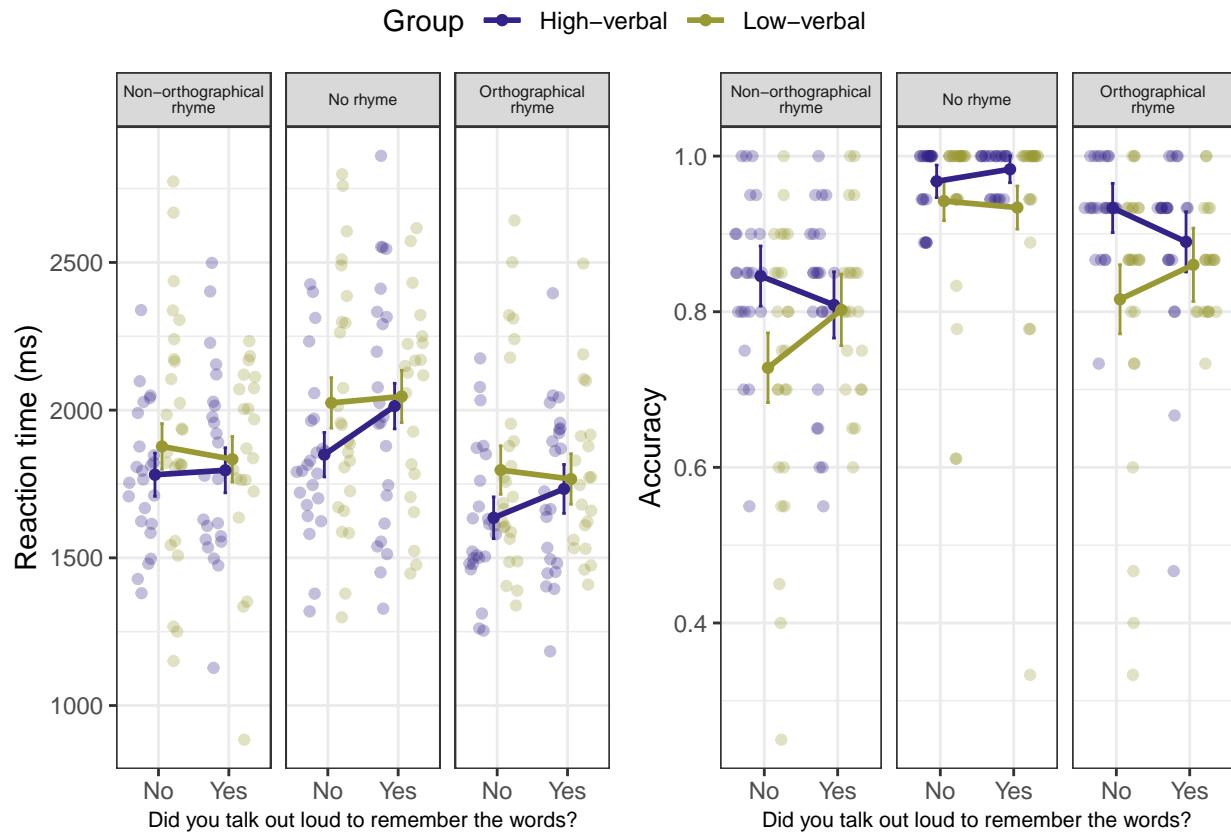
*Figure 9.* Reaction time and accuracy across groups by rhyme type. Error bars indicate 95 % confidence intervals.

<sup>371</sup> loud ( $\chi^2(1) = 0.01$ ,  $p = 0.91$ ). Nevertheless, the effect of doing so was interestingly different  
<sup>372</sup> for the two groups as can be seen in Figure 10.

<sup>373</sup> For both reaction time and accuracy, saying the words out loud diminished the  
<sup>374</sup> difference between the two groups. This suggests that this was the strategy that high-verbal  
<sup>375</sup> participants used in their heads - indeed, this was the most common strategy provided by  
<sup>376</sup> the participants (from both groups) who chose to answer the free answer about strategy.  
<sup>377</sup> There were no other notable strategies from the free answers.

### <sup>378</sup> 3.3 Task switching

<sup>379</sup> We excluded trials over 10 seconds (0.5 % of trials). We also recalculated the accuracy  
<sup>380</sup> measure so that any trial in the three switch conditions where participants in fact switched



*Figure 10.* Reaction time and accuracy by whether participants indicated that they had talked out loud to make the rhyme judgments. Error bars indicate 95 % confidence intervals.

381 between adding and subtracting counted as correct (as long as the arithmetic itself was also  
 382 correct). We did this to prevent a failure to switch once resulting in the remaining trials  
 383 counting as incorrect.

384       **3.3.1 Descriptive statistics: Task switching.** As can be seen from Table 4 and  
 385 Figure 11, accuracy was generally quite high in all conditions, and reaction times were  
 386 comparable across the two groups of participants.

387       **3.3.2 Statistical models: Task switching.** We conducted two linear mixed  
 388 models of verbal score and condition predicting reaction time and accuracy. Both models  
 389 included random intercepts for condition by participants, and the predictors were both  
 390 centered. The model predicting accuracy found that condition was a significant predictor  
 391 with an “increase” in condition being associated with lower accuracy ( $\beta = -0.58$ ; SE = 0.10;

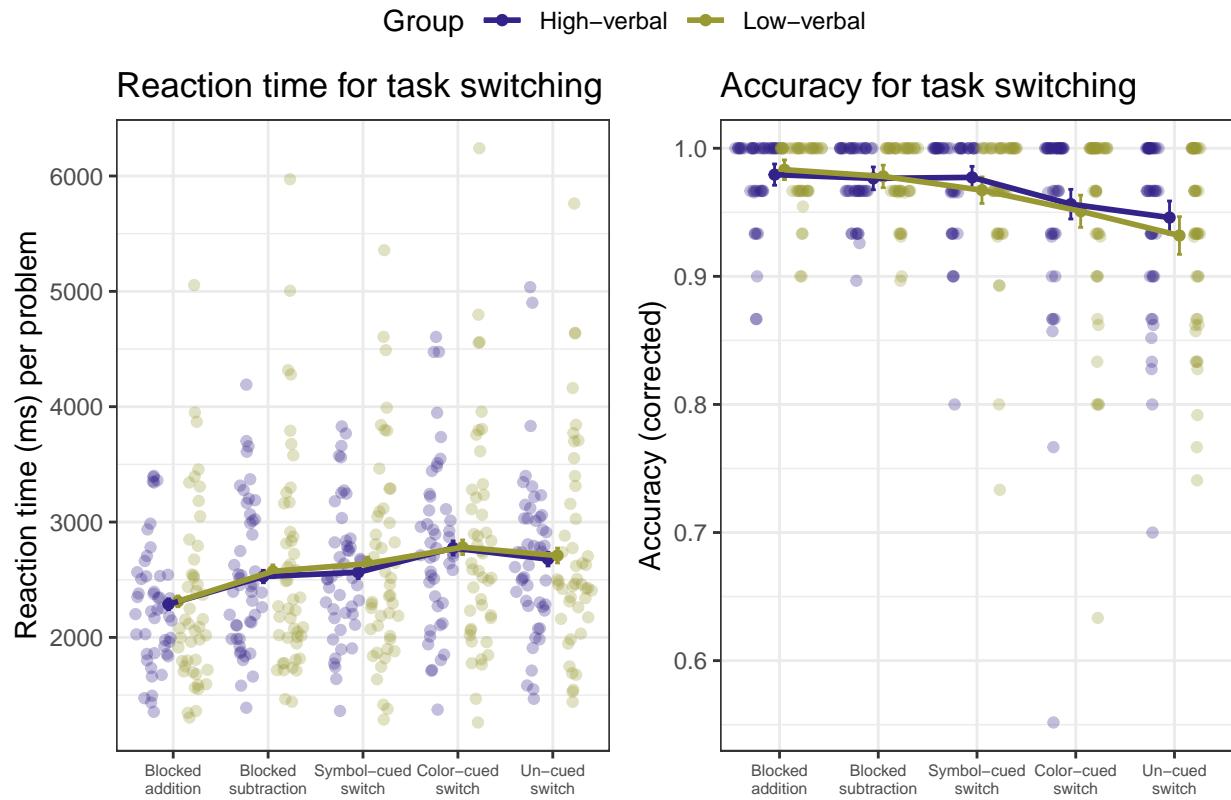


Figure 11. Reaction time and accuracy across conditions in the task switching experiment.

Error bars indicate 95 % confidence intervals.

<sup>392</sup>  $z = -5.50$ ;  $p < .001$ ). There was no effect of verbal score ( $p = 0.57$ ). However, there was a  
<sup>393</sup> marginally significant interaction effect with a higher verbal score diminishing the negative  
<sup>394</sup> effect of condition difficulty ( $\beta = 0.16$ ;  $SE = 0.08$ ;  $z = 1.87$ ;  $p = 0.06$ ). As for  
<sup>395</sup> log-transformed reaction time, there was a significant effect of condition difficulty with more  
<sup>396</sup> difficult conditions being associated with slower reaction times ( $\beta = 0.07$ ;  $SE = 0.02$ ;  $z =$   
<sup>397</sup>  $2.99$ ;  $p = 0.00$ ) but no effect of verbal score and no interaction effects (both  $p > 0.72$ ).

<sup>398</sup> If we calculate switch cost in a similar way to Emerson and Miyake (2003) (i.e., by  
<sup>399</sup> averaging problem solution time across simple addition and subtraction and then subtracting  
<sup>400</sup> the result from average problem solution time in the three switching conditions), we also do  
<sup>401</sup> not find any effects of verbal score on either reaction time ( $\beta = 3.46$ ;  $SE = 30.99$ ;  $t = 0.11$ ;  $p$   
<sup>402</sup>  $= 0.91$ ) or accuracy ( $\beta = -0.01$ ;  $SE = 0.00$ ;  $t = -1.21$ ;  $p = 0.23$ ). However, switch costs

403 differed between the three switching conditions with the uncued switch condition being  
 404 associated with higher reaction time switch costs ( $\beta = -67.29$ ; SE = 22.60;  $t = -2.98$ ;  $p =$   
 405 0.00) and higher accuracy switch costs ( $\beta = -0.02$ ; SE = 0.00;  $t = -4.42$ ;  $p < .001$ ).

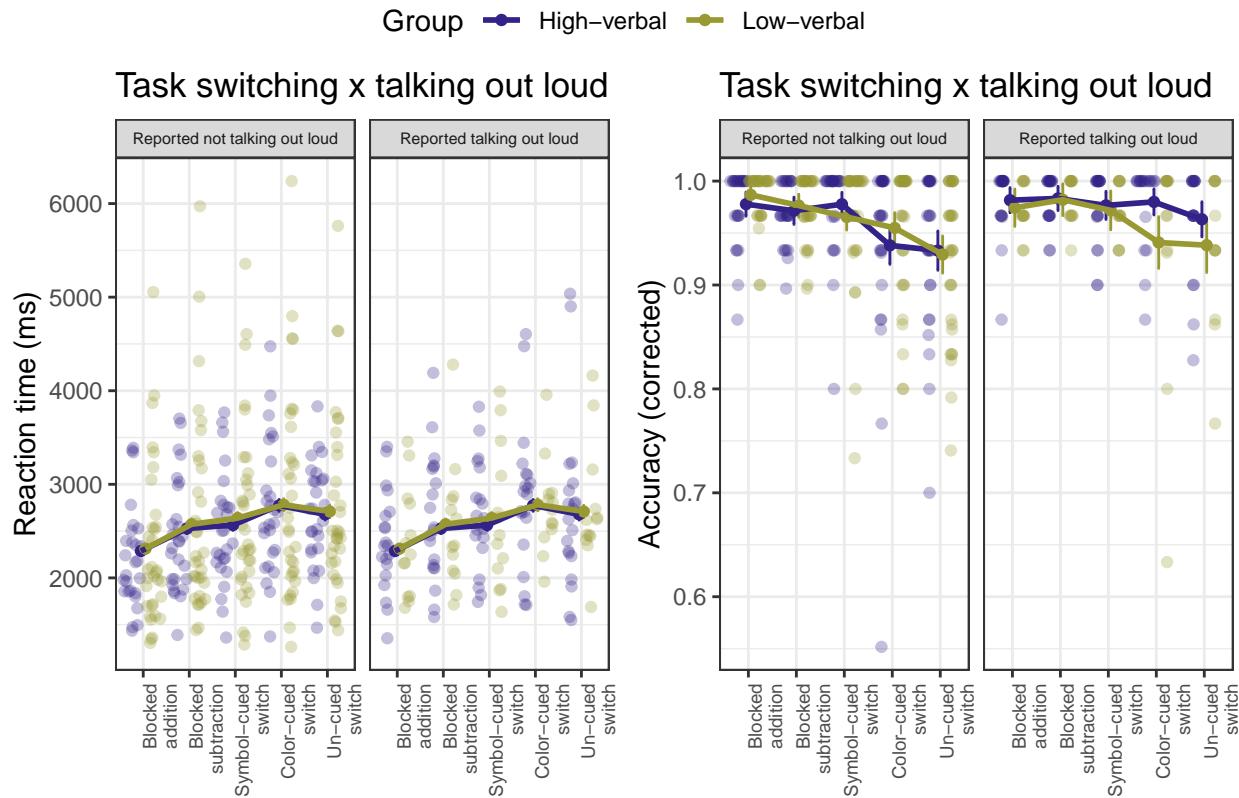


Figure 12. Accuracy in the task switching experiment by whether participants reported talking out loud to remember the correct rule or not.

406       **3.3.3 Strategies: Task switching.** We once again examined differences  
 407 associated with talking out loud, despite the fact that there were no general differences in  
 408 performance between the two groups. A chi-squared test showed that there was no  
 409 significant difference between how many high-verbal participants (20 out of 47) and how  
 410 many low-verbal participants (13 out of 46) reported that they had talked to themselves out  
 411 loud during the task ( $\chi^2(1) = 1.50$ ,  $p = 0.22$ ). There were not any obvious differences  
 412 between the effects that talking out loud had on these two groups (see accuracy and reaction  
 413 time Figure 12).

414 In response to the free answer question in the task switching experiment, several of the  
415 high-verbal participants said that they had said the answers out loud to themselves but not  
416 the operation ('add', 'subtract'). One visualized a cartoon character wearing red and giving  
417 thumbs-up or wearing blue and giving thumbs-down, one used their own thumb to keep  
418 track, and one used their fingers to count. Participants from the low-verbal group did not  
419 report many specific strategies apart from a few saying the operation or result out loud - one  
420 reported that they had tapped their index finger to mean 'add' and their middle finger to  
421 mean 'subtract'.

422 **3.4 Same/different judgments**

423 We excluded trials above 5 seconds (0.7 %) and below 200 ms (0.07 %).

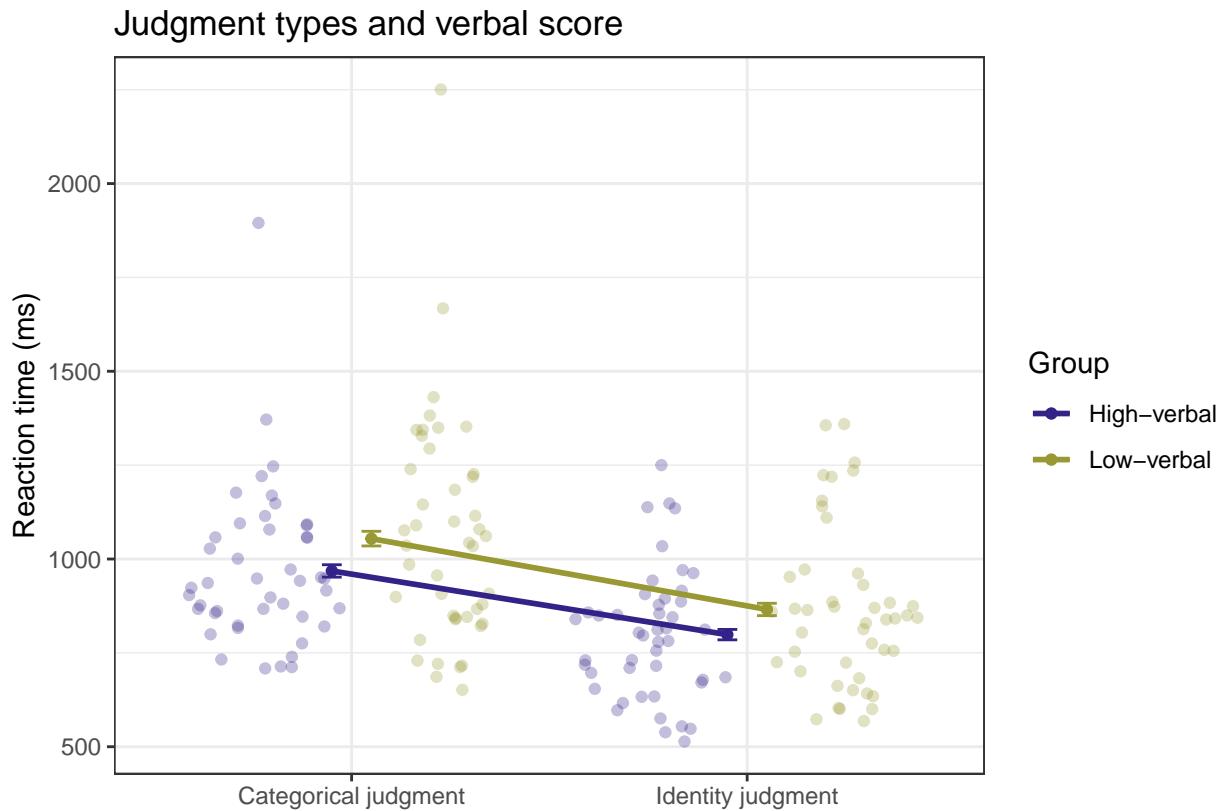
424 **3.4.1 Descriptive statistics by group: Same/different judgments.**

425 Generally, participants made the correct judgment on 95.53 % of trials. This did not differ  
426 between the high-verbal (95.58 %) and the low-verbal group (95.48 %). In subsequent  
427 analyses and plots, we only include correct trials. See Figure 13 for reaction times between  
428 the high-verbal and low-verbal groups for category judgments ('do these two animals belong  
429 to the same category?') or identity judgments ('are these two animals identical?').

430 **3.4.2 Statistical models: Same/different judgments.** We conducted a linear

431 mixed model of verbal score and judgment type predicting log-transformed reaction time  
432 including random intercepts per judgment type by participant. This model indicated  
433 significant main effect of judgment type and a marginally significant effect of verbal score.  
434 Identity judgments were faster than category judgments ( $\beta = -0.10$ ; SE = 0.01;  $t = -11.38$ ;  $p < .001$ ), and a higher verbal score was marginally associated with faster reaction times ( $\beta = -0.04$ ; SE = 0.02;  $t = -1.80$ ;  $p = 0.08$ ).

437 The key test for this experiment was whether the two groups behaved differently when  
438 giving correct 'DIFFERENT' responses on identity trials when the two images belonged to  
439 the same category. That is, we expected high-verbal participants to be more susceptible to



*Figure 13.* Reaction time in response to category or identity judgments. Error bars indicate 95 % confidence intervals.

<sup>440</sup> interference from a same-category distractor. See Figure 14.

<sup>441</sup> A linear mixed model of log-transformed reaction time with verbal score and category  
<sup>442</sup> membership of the distractor as predictors, including random intercepts per category  
<sup>443</sup> membership by participant, provided evidence that high-verbal participants were not  
<sup>444</sup> particularly affected by the within-category interference (interaction effect:  $p = 0.95$ ).  
<sup>445</sup> However, there was a significant main effect of category membership of the distractor with  
<sup>446</sup> within-category distractors being associated with slower reaction times ( $\beta = -0.04$ ; SE =  
<sup>447</sup> 0.01;  $t = -7.71$ ;  $p < .001$ ).

<sup>448</sup> **3.4.3 Additional analyses: Same/different judgments.** We also checked  
<sup>449</sup> whether the kind of animal made a difference on a within-category distractor trial. See  
<sup>450</sup> Figure 15.

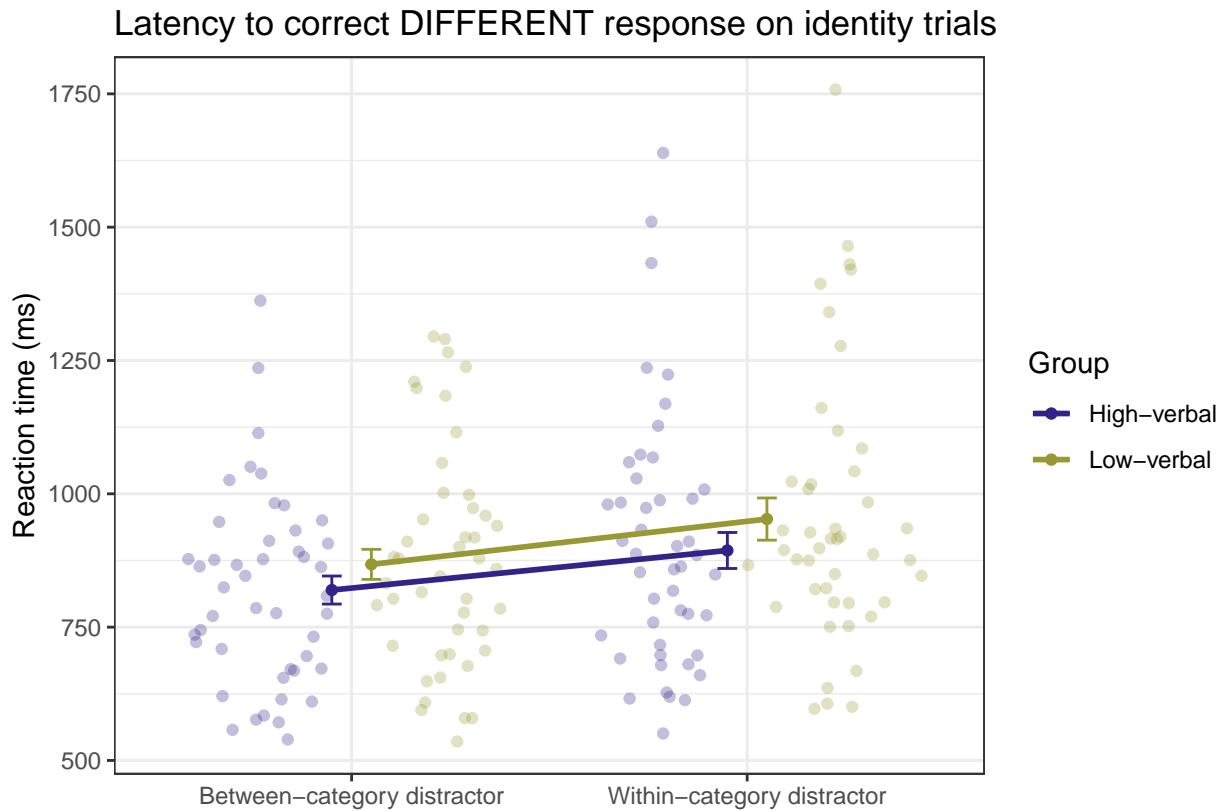


Figure 14. Reaction time on identity trials where the correct response was ‘DIFFERENT’ either because the two silhouettes were from different categories or different images from the same category. Error bars indicate 95 % confidence intervals.

451 A linear mixed model of log-transformed reaction times with verbal score and animal

452 pair (dog-dog or cat-cat) as predictors, including random intercepts per animal pair by

453 participant, provided evidence that dog-dog trials were faster than cat-cat trials ( $\beta = -0.03$ ;

454 SE = 0.01;  $t = -3.85$ ;  $p < .001$ ). However, this effect of animal pair was less strong when

455 verbal score was higher as indicated by a significant interaction effect between verbal score

456 and animal pair ( $\beta = 0.03$ ; SE = 0.01;  $t = 3.11$ ;  $p = 0.00$ ).

457 **3.4.4 Strategies: Same/different judgments.** In this experiment, most

458 participants said that they had no particular strategy. However, eight of the high-verbal

459 participants and one of the low-verbal participants explicitly mentioned something to do

460 with verbalizing the problems (e.g. ‘In my head I said “same” or “different” before I pressed

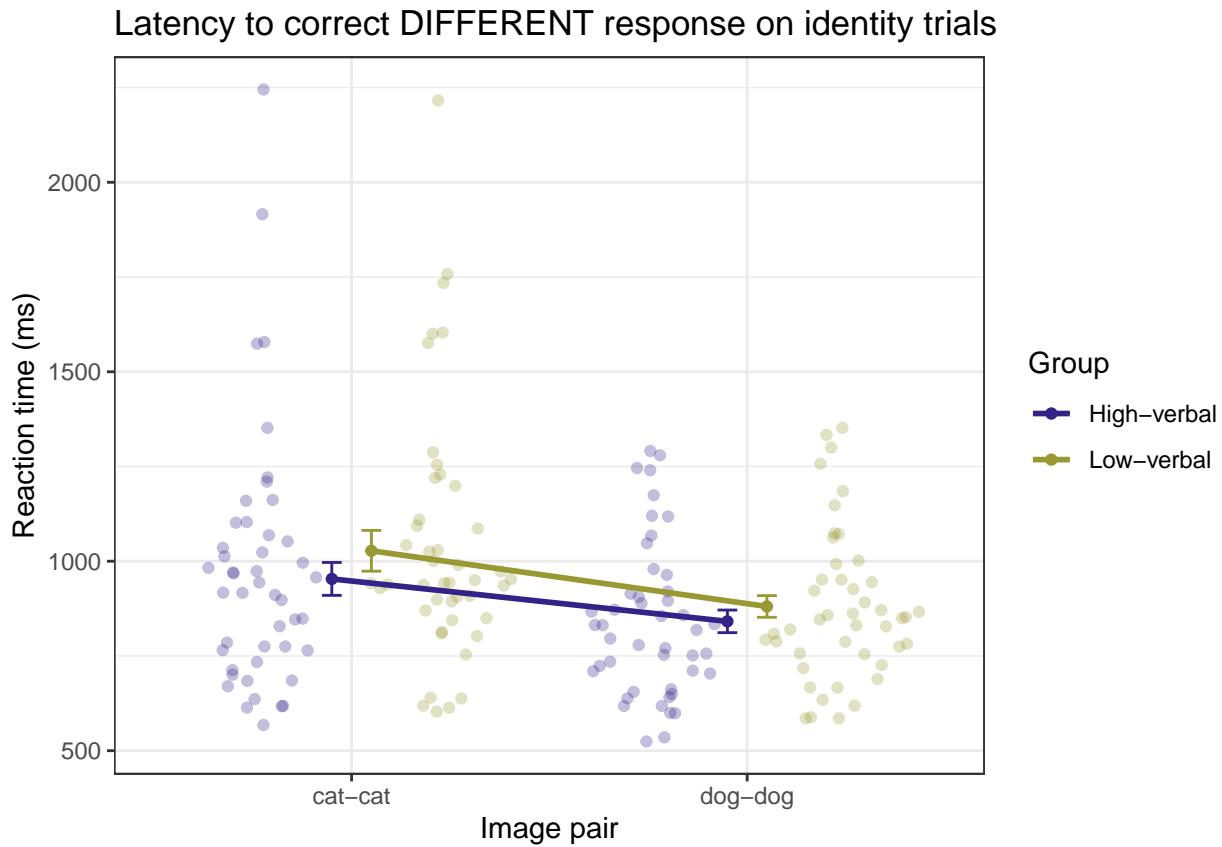


Figure 15. Reaction time in response to cat-cat or dog-dog judgments on identity trials when the correct response was ‘DIFFERENT’ (i.e., one cat and another cat or one dog and another dog). Error bars indicate 95 % confidence intervals.

<sup>461</sup> the arrow key.’)

### <sup>462</sup> 3.5 Intertask correlations

<sup>463</sup> We were interested in how performance on the different tasks correlated with each  
<sup>464</sup> other and whether these correlations were different for the two groups.

<sup>465</sup> **3.5.1 Overall intertask correlations.** See Figure 16. Generally, different  
<sup>466</sup> performance measures correlated within the same experiment. Interestingly, reaction times  
<sup>467</sup> on rhyming were negatively correlated with verbal working memory score suggesting some  
<sup>468</sup> working memory involvement in the rhyming task. Accuracy on uncued switch trials in the  
<sup>469</sup> task switching experiment was also positively correlated with accuracy on verbal working

470 memory and rhyming.

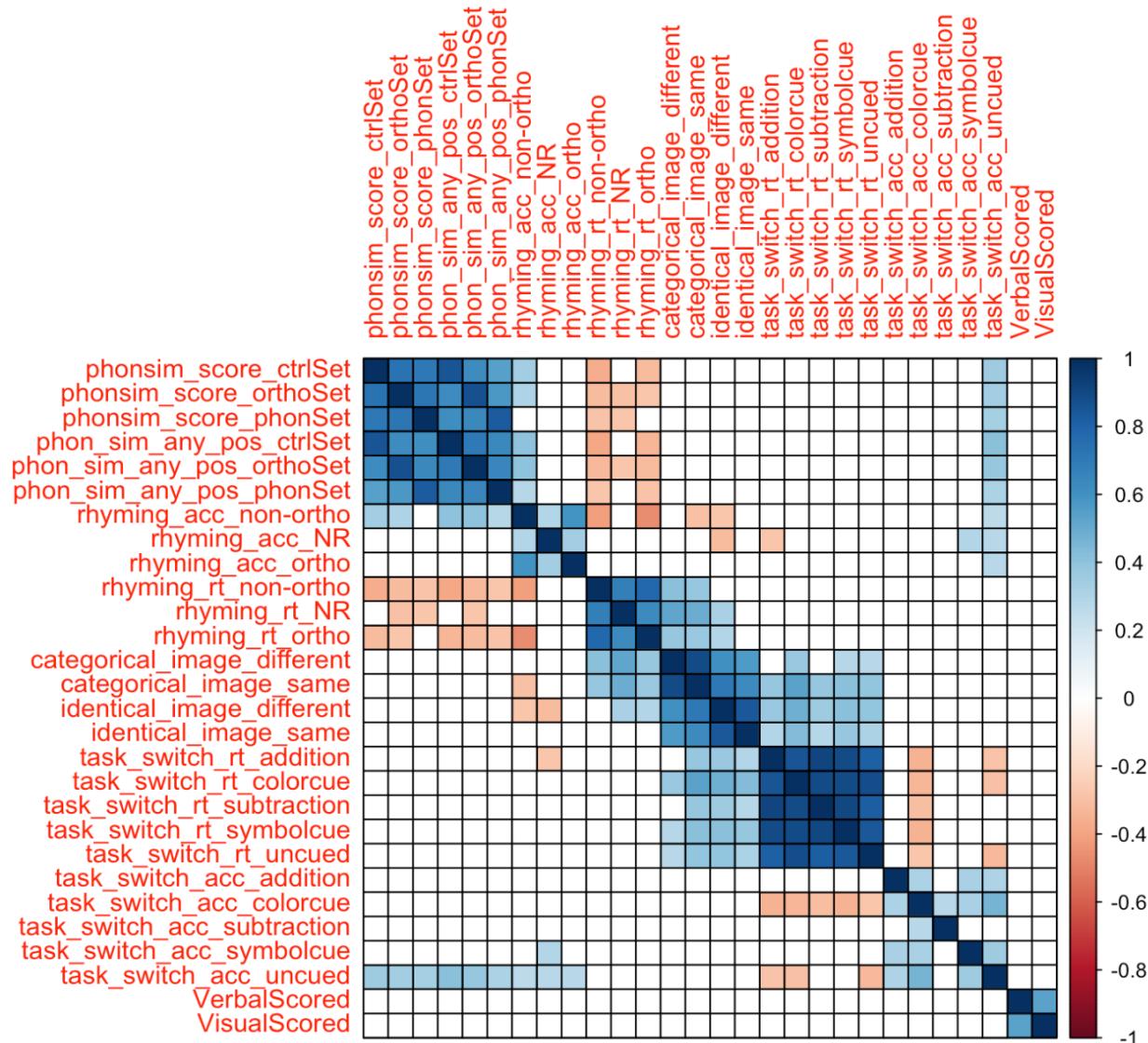
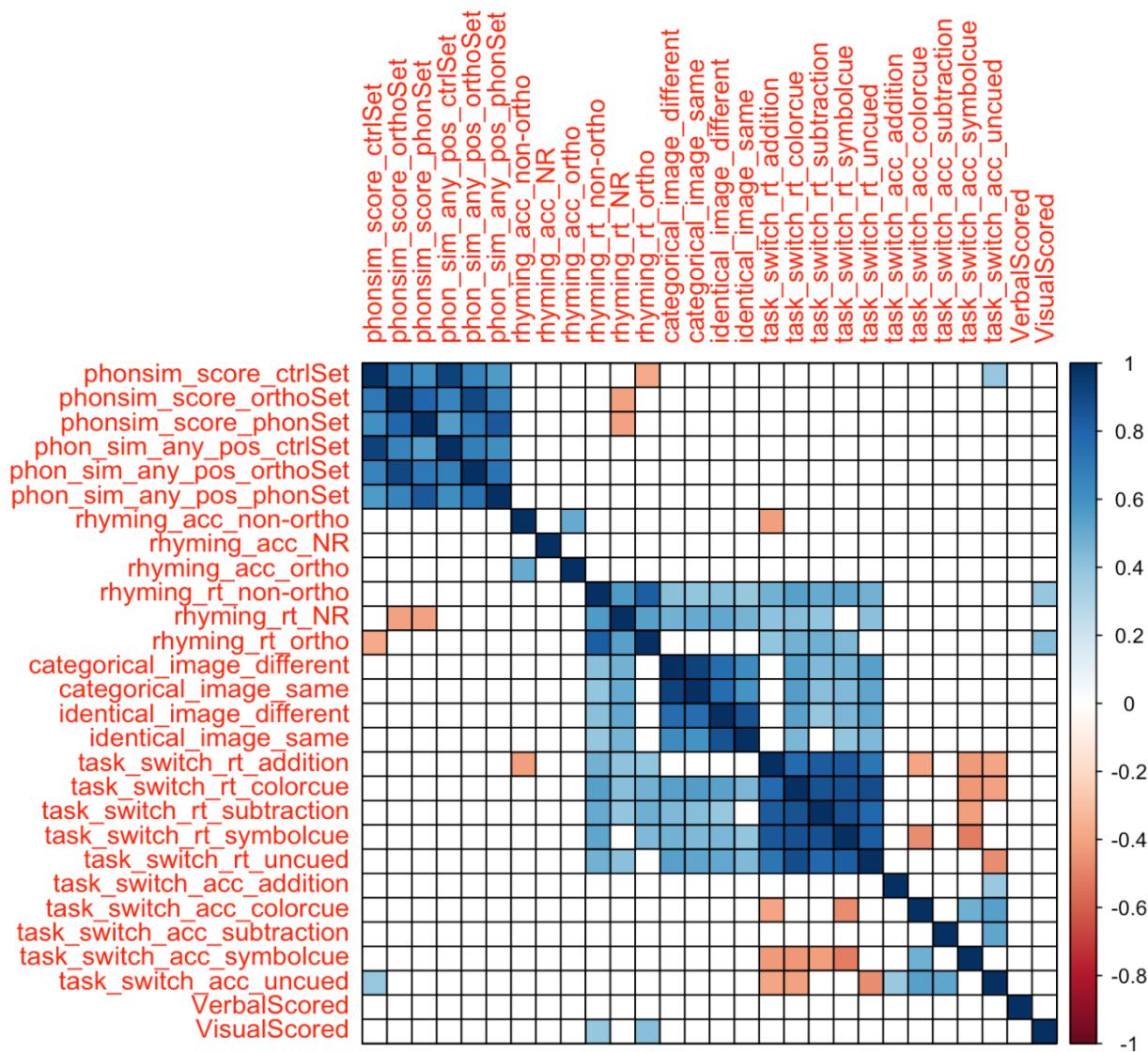


Figure 16. Intertask correlations in the total sample of high-verbal and low-verbal participants.

Colored squares represent significant correlations at  $p < .01$ .

471        3.5.2 Intertask correlations for the *high-verbal group*. See Figure 17. For  
 472 high-verbal participants, reaction times on rhyming, category judgments, and task switching  
 473 were positively correlated, suggesting that these relied on similar mechanisms for this group.



*Figure 17.* Intertask correlations within the high-verbal group. Colored squares represent significant correlations at  $p < .01$ .

### 474 3.5.3 Intertask correlations for the *low-verbal group*. See Figure 18. For

475 low-verbal participants, there was no such widespread correlation between rhyming, category  
 476 judgments, and task switching reaction times. However, they showed a positive correlation  
 477 between accuracy on uncued switch trials and (position-indifferent) verbal working memory  
 478 and rhyming accuracy.

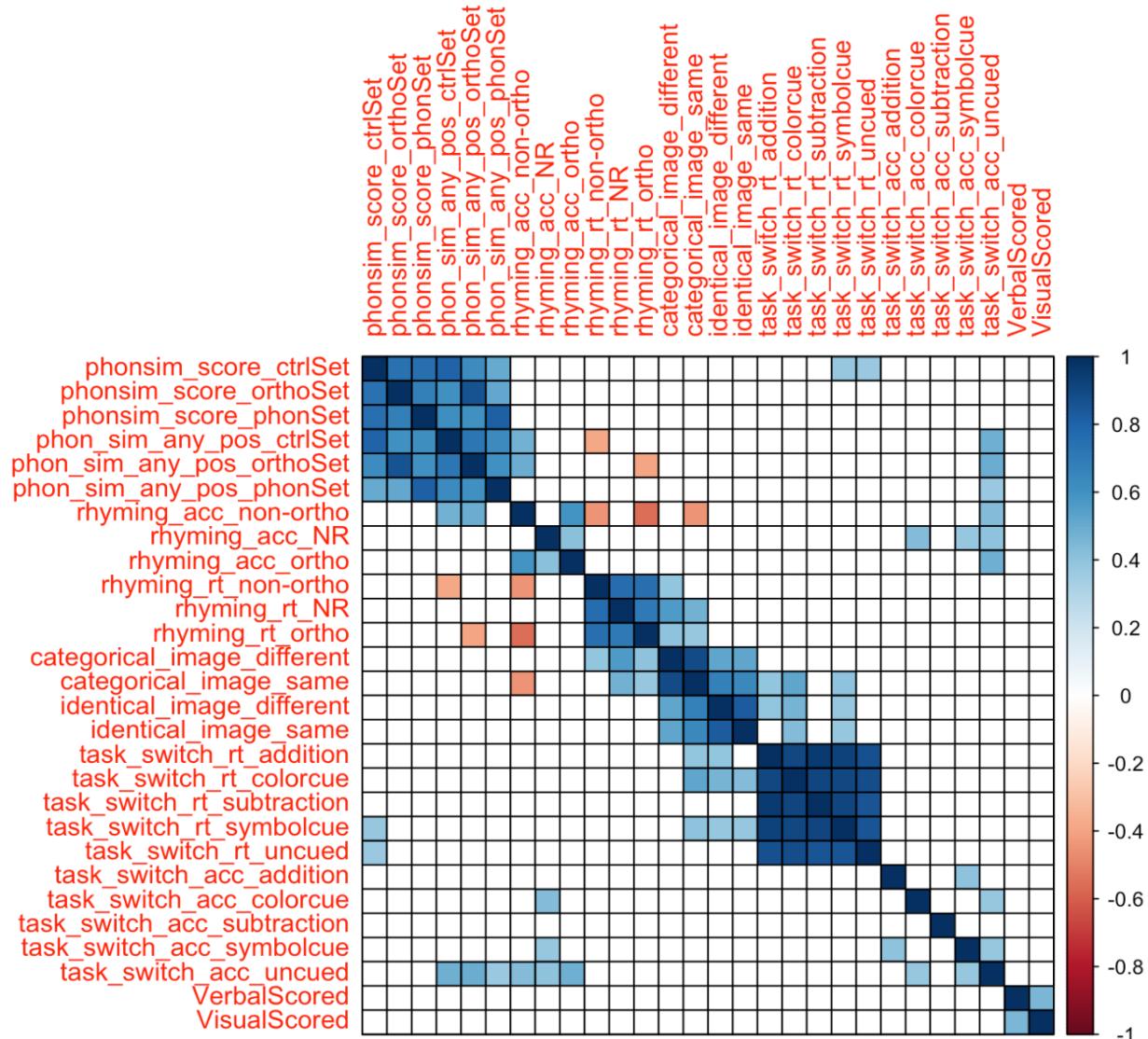


Figure 18. Intertask correlations within the low-verbal group. Colored squares represent significant correlations at  $p < .01$ .

### 479 3.5.4 Difference between task correlations in *low-verbal group* and

480 ***high-verbal group.*** See Figure 19. Note that colored squares do not have any  
 481 relationship with statistical significance on this plot. Blue simply means that when the  
 482 correlation coefficient for the high-verbal group was subtracted from the correlation  
 483 coefficient for the low-verbal group, the result was positive. Similarly, red-orange means that  
 484 when the correlation coefficient for the high-verbal group was subtracted from the correlation

485 coefficient for the low-verbal group, the result was negative.

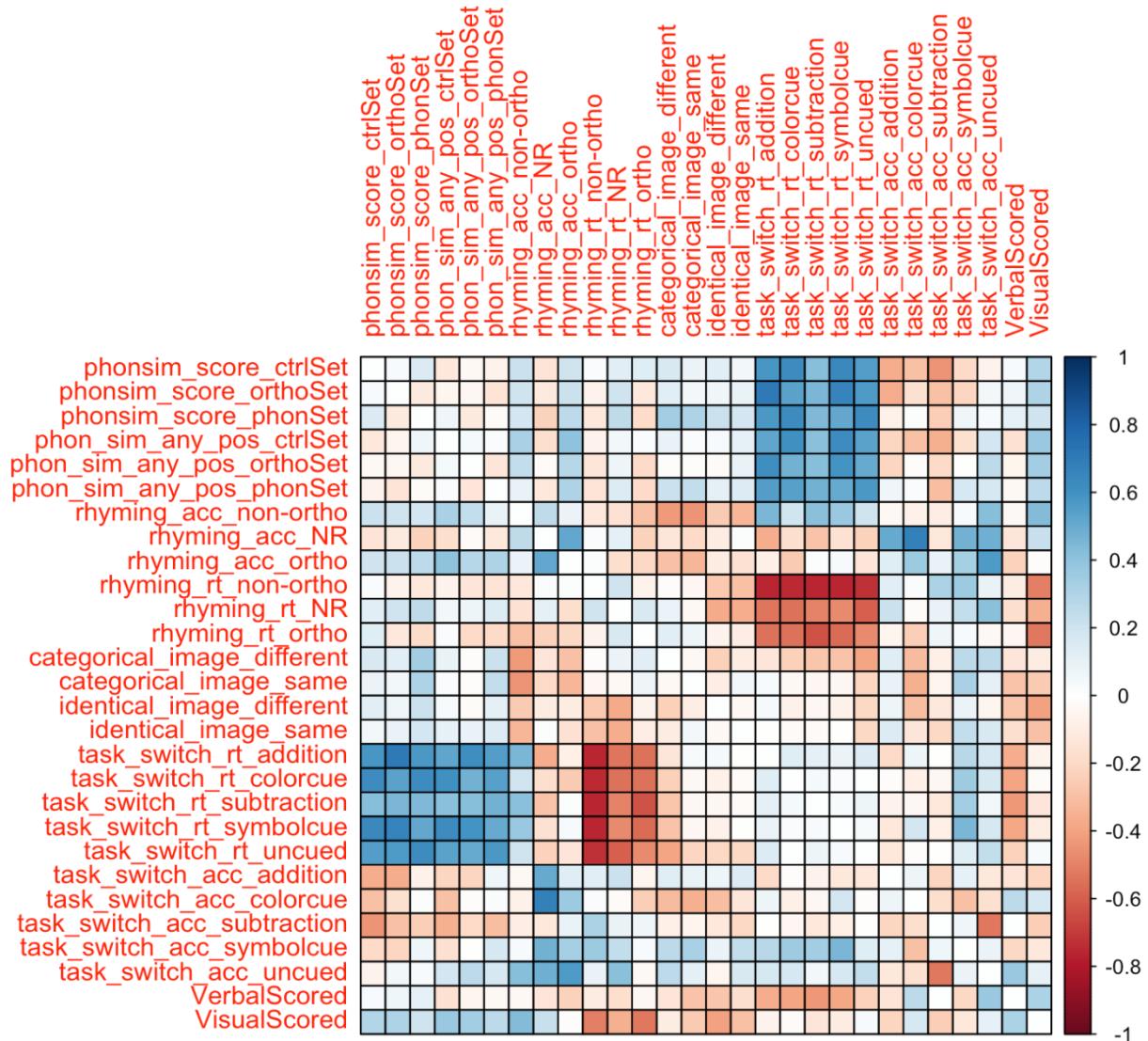


Figure 19. Intertask correlations in the high-verbal group subtracted from intertask correlations in the low-verbal group.

### 486 3.6 Summary of behavioral findings

487 In the phonological similarity experiment, the high-verbal group performed better at  
 488 both position-specific recall and position-indifferent recall. The difference between the two  
 489 groups was diminished for the orthographically similar set when considering

490 position-indifferent recall. Interestingly, the differences between the two groups was  
491 diminished if participants reported talking out loud to remember the words.

492 In the rhyming experiment, high-verbal participants performed significantly better  
493 than low-verbal participants. Once again, the differences in both accuracy and reaction time  
494 between the two groups was eliminated when participants reported naming the pictures out  
495 loud.

496 Participants who scored lower on verbal representations were slower at making  
497 category-based judgments in the same-different experiment. We expected high-verbal  
498 participants to show a more marked effect of being distracted by category membership on  
499 identity judgment trials (e.g. being slower at correctly responding ‘different’ to two pictures  
500 of different cats), but this did not appear to be the case.

501 There were no notable differences between the two groups in the task switching  
502 experiment.

503 **3.7 Questionnaire measures**

504 **For some strange reason, we do not have questionnaire data from**  
505 **A3KVKK1XLBTSN3.** We will retain their data from the four behavioral experiments  
506 and here report questionnaire data from 47 high-verbal and 45 low-verbal participants.

507 For most of our custom questions, there were notable differences in how participants  
508 from the two groups responded (see Figure 20). The questions with the clearest differences  
509 concerned rehearsing and revising conversations where the high-verbal group reported doing  
510 so much more often than the low-verbal group did.

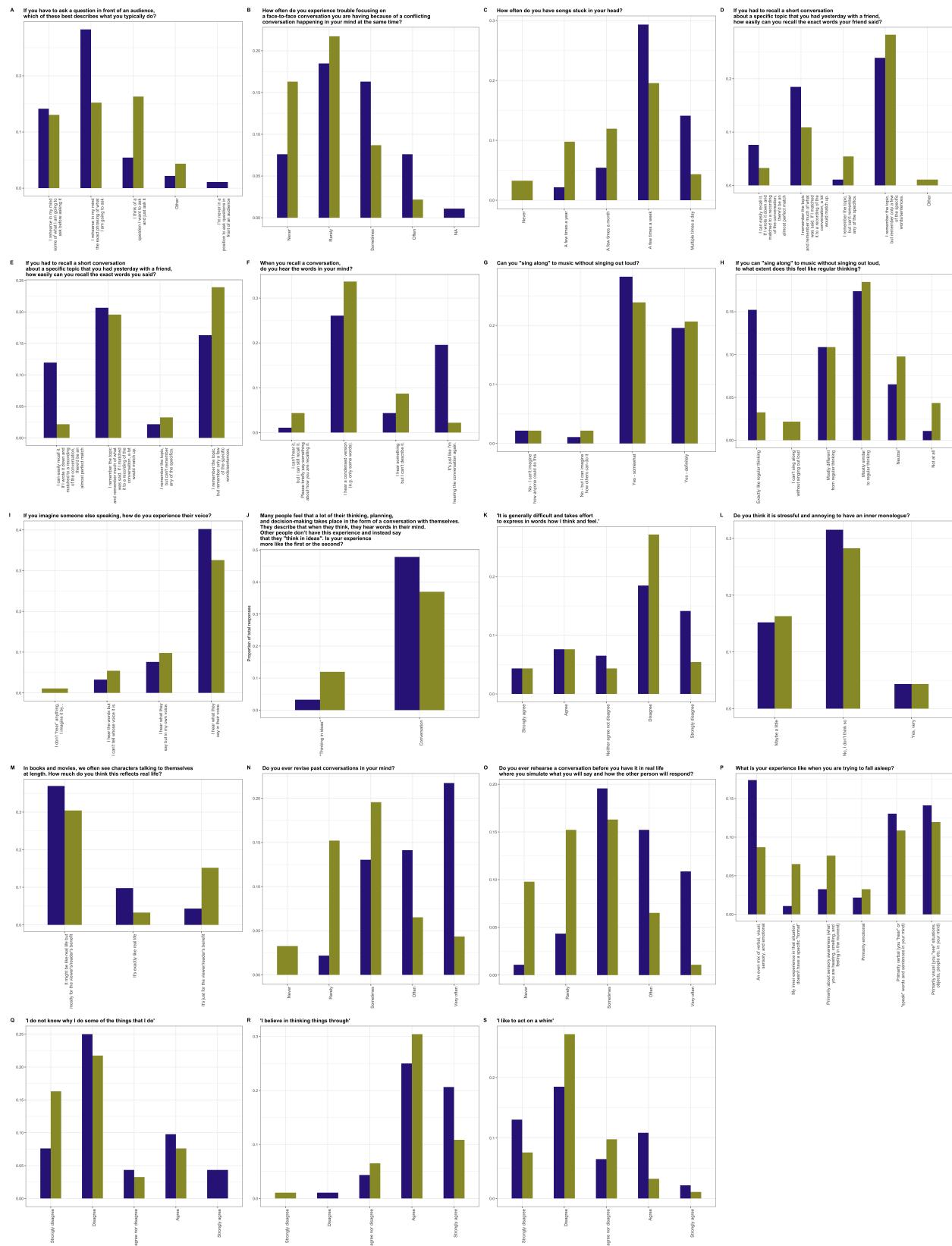


Figure 20. Grouped bar plots of our custom questions. Dark blue represents the high-verbal group and dark yellow represents the low-verbal group.

511

#### 4 Discussion

512 In this exploratory study, we found significant behavioral differences between  
513 participants who experience a lot of verbal representations and participants who do not.  
514 High-verbal participants were more accurate at making rhyme judgments from images, faster  
515 at making both category and identity judgments for simple visual stimuli, and showed better  
516 verbal working memory performance. The only experiment where we found no performance  
517 differences between the two groups was the task switching experiment. Interestingly, in both  
518 the rhyming experiment and the verbal working memory experiment, performance differences  
519 between the two groups disappeared when participants reported talking out loud to solve the  
520 problems. In terms of our questionnaire, the differences were as expected: High-verbal  
521 participants were more likely to say that they rehearse or revise conversations, and that their  
522 experience of thinking about or recalling conversations is much like hearing every word.  
523 They also reported having songs stuck in their head more often and were less likely to say  
524 that the way inner speech is represented in media is for the reader's/viewer's benefit.

525 Having found evidence that a diminished experience of inner speech has behavioral  
526 consequences, we would like to propose a name for the phenomenon: *anendophasia*. We  
527 came up with this in collaboration with people on Reddit and in real life who experience no  
528 inner speech. Many of them objected to words implying silent or quiet minds as they  
529 experience their minds as quite busy. Similarly, they did not find words implying lack of  
530 mental language appropriate as they experience words/concepts but not speech. We believe  
531 *anendophasia* is suitable because it includes components that people are likely to already be  
532 familiar with ('-phasia' as something to do with speech) and targets specifically speech and  
533 not words per se. Furthermore, the term endophasia already exists as a term for inner  
534 speech (Bergounioux, 2001; Loevenbruck et al., 2018). We do not believe a lack of inner  
535 speech could simply be subsumed under the umbrella term 'aphantasia' for lack of mental  
536 imagery (Monzel, Mitchell, Macpherson, Pearson, & Zeman, 2022) because inner speech is

537 both auditory and articulatory in nature (whether it is better termed ‘inner hearing’ or  
538 ‘inner speaking’ is also subject to debate) and because the linguistic properties of inner  
539 speech are not reducible to phonological properties. For these reasons, we also do not believe  
540 the previously proposed term *anauralia* is appropriate (Hinwar & Lambert, 2021).

541 **4.1 What have we learned about people with anendophasia?**

542 Most importantly, it appears that people are right when they report that their  
543 experience rarely takes a verbal format. This is not self-evident as Descriptive Experience  
544 Sampling has sometimes found that people can be quite mistaken about the format of their  
545 inner experience (Heavey & Hurlburt, 2008; Russell T. Hurlburt, 2011; Russell T. Hurlburt,  
546 Heavey, & Kelsey, 2013). On our tasks, participants with anendophasia seemed less able to  
547 use internal verbalization to increase performance. This is especially interesting as the  
548 questions that are related to the verbal factor on the Internal Representations Questionnaire  
549 (Roebuck & Lupyan, 2020) and which we used for participant selection are about the format  
550 of spontaneous thought (e.g., ‘I think about problems in my mind in the form of a  
551 conversation with myself’ and ‘If I am walking somewhere by myself, I often have a silent  
552 conversation with myself’). There is some evidence that spontaneously occurring inner  
553 speech and experiment-elicited inner speech are not necessarily comparable and have  
554 different neural substrates (Russell T. Hurlburt, Alderson-Day, Kühn, & Fernyhough, 2016)  
555 which makes it remarkable that our participants’ reports of spontaneous inner speech seem  
556 related to their ability to use internal verbalization and verbal working memory. It is also  
557 interesting that performance was in many cases related to verbal score as a continuous factor  
558 which indicates that anendophasia is not an all-or-nothing phenomenon, much like  
559 aphantasia is not (Dance et al., 2022).

560 Regarding the parallels with aphantasia, it is important to note that the analogy can  
561 only take us so far. Given the findings from the present study, it seems unlikely that people  
562 with anendophasia are completely unable to verbalize internally like some people with

aphantasia are completely unable to visualize. If you ask somebody with anendophasia to read something out loud in their head, for example, they can do it. What seems instead to be the case is that they do not use or only rarely use inner speech spontaneously in everyday life to plan, solve problems, and rehearse conversations (Perrone-Bertolotti et al., 2014). A question remains of how thinking is experienced if not verbally. As discussed in the Introduction, there is anecdotal evidence that some people with anendophasia have very visual experience and others “think in ideas”. This largely corresponds to what we found when we asked participants whether their thinking was more like a conversation or more like “thinking in ideas”. One participant described their thinking as ‘I really do think in concepts rather than forming words in my head’ and another reported ‘I visualize what I am trying to do or plan and act accordingly’. We are not much closer to understanding what “thinking in ideas” means but note that it is probably related to the unsymbolized thinking of Hurlburt and colleagues.

We did find evidence that using other strategies than internal verbalization could diminish the performance differences between our two groups. This was clearest when we examined whether participants reported talking out loud to solve the problems or not. In both the verbal working memory experiment and in the rhyme judgment experiment, performance differences disappeared when participants reported talking out loud. This suggests that participants without anendophasia were already using verbalization strategies internally. One particularly interesting example comes from orthographically similar words in the verbal working memory experiment ('rough', 'cough', 'through', 'dough', 'bough'). Many participants with anendophasia reported a strategy of remembering just the first letters of the words once they were familiar with the set, thus reducing the load on verbal working memory. This is likely to be the reason why there was no difference in performance between the two groups for this word set. Another interesting case is the finding that the two groups did not differ in either reaction time or performance on the task switching experiment. This suggests that while the inner voice can be used as a behavioral self-cue,

590 other and equally effective strategies may be available.

## 591 4.2 Implications for the relationship between language and cognition

592 In our view, the most interesting question that the present paper can help address is  
593 this: If there are indeed people who have no inner voice and do not use language to think,  
594 and if these people behave more or less like everyone else (hence the masked individual  
595 differences), does this mean that the connection between language and thought could in fact  
596 be quite weak? We saw that participants with anendophasia performed worse on tasks  
597 specifically requiring internal verbalization (verbal working memory and rhyme judgments)  
598 and were also slower on both category and identity judgments. However, they did not differ  
599 from our high-verbal control group in performance on a task switching paradigm – this is  
600 surprising because verbal interference studies have previously suggested that people use their  
601 inner voice to cue themselves on the relevant task rule (plus or minus). Our interpretation of  
602 our findings is that inner speech can be recruited for self-cuing but that self-cuing can be as  
603 effectively achieved through other means as well.

604 Taking the findings from our experiments together, it seems unlikely that inner speech  
605 plays a constitutive role in cognition, although we allow that this may still be the case for  
606 some kinds of cognition (e.g., metacognition and self-reflection, see Morin (2018) and  
607 Bermúdez (2007)). Inner speech is instead plausibly recruited as a tool by some people for  
608 some tasks which is a much weaker connection than has previously been suggested. When  
609 inner speech is recruited for tasks, it provides format that supports linear order (as seen in  
610 our verbal working memory experiment) and more readily available categorization (as seen in  
611 our same-different experiment). Our study is merely a preliminary exploration, and there are  
612 many other kinds of thinking that should be tested as a next step. Candidates include  
613 abstraction, relational thinking, metacognition, and social cognition, all of which have been  
614 hypothesized to rely on language to some extent.

### 615 4.3 Limitations of the present study

616 It may be the case that the performance differences we found were driven by some  
617 other factor than how verbally represented participants' thoughts were. For example,  
618 high-verbal participants could be generally more motivated or conscientious than low-verbal  
619 participants. We believe that this explanation is unlikely, however, since we saw examples of  
620 specific conditions where there were no differences between the two groups (e.g., responses to  
621 the no-rhyme pairs in the rhyme judgment experiment, responses to the orthographically  
622 similar set in the verbal working memory experiment, and all conditions in the task  
623 switching experiment).

### 624 4.4 Future directions

625 The present work is exploratory and presents a few examples of how we might  
626 empirically test the consequences of having or not having an inner voice. Some of the most  
627 interesting avenues forward concern alternative strategies as well as how a lack of inner  
628 speech might influence other things than behavior such as mental health and social cognition.

629 Just as in aphantasia, it could be the case that individual differences in inner speech  
630 remain largely undiscovered because people use alternative but equally efficient strategies for  
631 solving problems. We saw this for example in the aphantasia study where participants with  
632 aphantasia performed equally well but had a different response profile (Keogh et al., 2021).  
633 We see some indications in our present study as well with the intertask correlations being  
634 different in the low- and high-verbal groups. This should be explored in future studies – the  
635 optimal way would be to design an experiment where different strategies would show  
636 different behavioral profiles but this is quite difficult.

637 Many areas of research and therapy (most notably Cognitive Behavior Therapy and  
638 many mindfulness practices) rely heavily on inner speech and are difficult to tailor to people  
639 who find this unnatural. A natural first step would be to ask people how much they

640 spontaneously use their inner voice in everyday life and take this into account when  
641 designing interventions.

642 The largest difference between the two groups in the questionnaire was that  
643 high-verbal participants spent more time rehearsing conversations which makes us wonder  
644 what kind of consequences this might have. Would we expect high- verbal people to be  
645 somehow “better” at conversations? Or maybe worse because they over-rehearse? It seems  
646 that inner speech is very linked to social interactions so in future studies we would like to  
647 assess social cognitive abilities in populations with and without habitual inner speech. This  
648 could for example be with an adult version of the Faux Pas test (e.g., Baron-Cohen,  
649 ORiordan, Stone, Jones, & Plaisted, 1999; Thiébaut et al., 2015).

650 **5 Conclusion**

651 In four exploratory behavioral experiments, we found significant differences between  
652 participants whose inner experience is verbal to a large extent and participants whose inner  
653 experience is verbal to a much lesser extent. We propose a term for the experience of the  
654 latter group: anendophasia. Participants with anendophasia were worse at making rhyme  
655 judgments in response to images, remembering words in order, and slower at making both  
656 category and identity judgments of simple visual stimuli. They were not, however, either  
657 slower or less accurate at a task switching experiment requiring them to switch between  
658 addition and subtraction problems. Taken together, our experiments suggest that there are  
659 real behavioral consequences of experiencing less or more inner speech, and that these  
660 differences may usually be masked because people with anendophasia use alternative  
661 strategies. Our findings have important implications for theories about the relationship  
662 between inner speech and cognition.

663

## 6 References

- 664 Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions,  
665 phenomenology, and neurobiology. *Psychological Bulletin*, 141(5), 931.
- 666 Alderson-Day, B., Mitrenga, K., Wilkinson, S., McCarthy-Jones, S., & Fernyhough, C.  
667 (2018). The varieties of inner speech questionnaire—revised (VISQ-r): Replicating and  
668 refining links between inner speech and psychopathology. *Consciousness and Cognition*,  
669 65, 48–58.
- 670 Aust, F., & Barth, M. (2022). *papaja: Prepare reproducible APA journal articles with R*  
671 *Markdown*. Retrieved from <https://github.com/crsh/papaja>
- 672 Baddeley, A. (1966). Short-term memory for word sequences as a function of acoustic,  
673 semantic and formal similarity. *Quarterly Journal of Experimental Psychology*, 18(4),  
674 362–365.
- 675 Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action:  
676 Evidence from task switching. *Journal of Experimental Psychology: General*, 130(4), 641.
- 677 Baron-Cohen, S., ORiordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). *Journal of*  
678 *Autism and Developmental Disorders*, 29(5), 407–418.  
679 <https://doi.org/10.1023/a:1023035012436>
- 680 Barth, M. (2022). *tinylabes: Lightweight variable labels*. Retrieved from  
681 <https://cran.r-project.org/package=tinylabes>
- 682 Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models  
683 using lme4. *Journal of Statistical Software*, 67(1), 1–48.  
684 <https://doi.org/10.18637/jss.v067.i01>
- 685 Bates, D., & Maechler, M. (2021). *Matrix: Sparse and dense matrix classes and methods*.  
686 Retrieved from <https://CRAN.R-project.org/package=Matrix>
- 687 Bergounioux, G. (2001). Endophasie et linguistique [décomptes, quotes et squelette]. *Langue*  
688 *Française*, 132, 106–124.
- 689 Bermúdez, J. L. (2007). *Thinking without words*. Oxford University Press.

- 690 Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*,  
691 25(6), 657–674.
- 692 Carruthers, P. (2009). Mindreading underlies metacognition. *Behavioral and Brain Sciences*,  
693 32(2), 164–182.
- 694 Chella, A., & Pipitone, A. (2020). A cognitive architecture for inner speech. *Cognitive*  
695 *Systems Research*, 59, 287–292.
- 696 Clark, A. (1998). *Language and thought: Interdisciplinary themes* (P. Carruthers & J.  
697 Boucher, Eds.). Cambridge University Press.
- 698 Cragg, L., & Nation, K. (2010). Language and the development of cognitive control. *Topics*  
699 *in Cognitive Science*, 2(4), 631–642.
- 700 Dance, C., Ipser, A., & Simner, J. (2022). The prevalence of aphantasia (imagery weakness)  
701 in the general population. *Consciousness and Cognition*, 97, 103243.
- 702 Dawes, A. J., Keogh, R., Andrillon, T., & Pearson, J. (2020). A cognitive profile of  
703 multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, 10(1),  
704 1–10.
- 705 De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in  
706 a web browser. *Behavior Research Methods*, 47(1), 1–12.
- 707 Dowle, M., & Srinivasan, A. (2021). *Data.table: Extension of ‘data.frame’*. Retrieved from  
708 <https://CRAN.R-project.org/package=data.table>
- 709 Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., &  
710 Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six  
711 european languages. *Quarterly Journal of Experimental Psychology*, 71(4), 808–816.
- 712 Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A  
713 dual-task investigation. *Journal of Memory and Language*, 48(1), 148–168.
- 714 Faw, B. (2009). Conflicting intuitions may be based on differing abilities: Evidence from  
715 mental imaging research. *Journal of Consciousness Studies*, 16(4), 45–68.
- 716 Felton, J. (2020). People with no internal monologue explain what it’s like in their head.

- 717 IFLScience. Retrieved from <https://www.iflscience.com/people-with-no-internal->  
718 monologue-explain-what-its-like-in-their-head-57739
- 719 Frankish, K. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente, Eds.).  
720 Oxford University Press.
- 721 Galton, F. (1880). Statistics of mental imagery. *Mind*, 5(19), 301–318.
- 722 Gauker, C. (2011). *Words and images: An essay on the origin of ideas*. Oxford University  
723 Press, Oxford.
- 724 Geva, S. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente, Eds.). Oxford  
725 University Press.
- 726 Geva, S., Bennett, S., Warburton, E. A., & Patterson, K. (2011). Discrepancy between inner  
727 and overt speech: Implications for post-stroke aphasia and normal language processing.  
728 *Aphasiology*, 25(3), 323–343.
- 729 Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in  
730 the right visual field but not the left. *Proceedings of the National Academy of Sciences*,  
731 103(2), 489–494.
- 732 Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2008). Support for lateralization of the  
733 whorf effect beyond the realm of color discrimination. *Brain and Language*, 105(2),  
734 91–98.
- 735 Harrell Jr, F. E., Charles Dupont, with contributions from, & others., many. (2021). *Hmisc:*  
736 *Harrell miscellaneous*. Retrieved from <https://CRAN.R-project.org/package=Hmisc>
- 737 Heavey, C. L., & Hurlburt, R. T. (2008). The phenomena of inner experience. *Consciousness  
738 and Cognition*, 17(3), 798–810.
- 739 Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from  
740 <https://CRAN.R-project.org/package=purrr>
- 741 Hinwar, R. P., & Lambert, A. J. (2021). Anauralia: The silent mind and its association with  
742 aphantasia. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.744213>
- 743 Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., et al.others.

- 744 (2007). Functional and morphometric brain dissociation between dyslexia and reading  
745 ability. *Proceedings of the National Academy of Sciences*, 104(10), 4234–4239.
- 746 Hurlburt, Russell T. (2011). *Investigating pristine inner experience*. Cambridge University  
747 Press. <https://doi.org/10.1017/cbo9780511842627>
- 748 Hurlburt, Russell T., & Akhter, S. A. (2006). The descriptive experience sampling method.  
749 *Phenomenology and the Cognitive Sciences*, 5(3), 271–301.
- 750 Hurlburt, Russell T., & Akhter, S. A. (2008). Unsymbolized thinking. *Consciousness and  
751 Cognition*, 17(4), 1364–1374.
- 752 Hurlburt, Russell T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the  
753 ecological validity of thinking on demand: Neural correlates of elicited vs. Spontaneously  
754 occurring inner speech. *PLoS One*, 11(2), e0147932.
- 755 Hurlburt, Russell T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of  
756 inner speaking. *Consciousness and Cognition*, 22(4), 1477–1494.  
757 <https://doi.org/10.1016/j.concog.2013.10.003>
- 758 Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance  
759 in aphantasia. *Cortex*, 105, 61–73.
- 760 Kassambara, A. (2020). *Ggpubr: 'ggplot2' based publication ready plots*. Retrieved from  
761 <https://CRAN.R-project.org/package=ggpubr>
- 762 Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia.  
763 *Cortex*, 105, 53–60.
- 764 Keogh, R., Wicken, M., & Pearson, J. (2021). Visual working memory in aphantasia:  
765 Retained accuracy and capacity with a different strategy. *Cortex*, 143, 237–253.
- 766 Kothe, E., Callegher, C. Z., Gambarota, F., Linkersdörfer, J., & Ling, M. (2021). *Trackdown:  
767 Collaborative writing and editing of r markdown (or sweave) documents in google drive*.  
768 <https://doi.org/10.5281/zenodo.5167320>
- 769 Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests  
770 in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.

- 771 https://doi.org/10.18637/jss.v082.i13
- 772 @KylePlantEmoji. (2020). Fun fact: Some people have an internal narrative and some don't.  
773 Twitter. Retrieved from https://twitter.com/KylePlantEmoji/status/1221713792913965  
774 061?s=20&t=JMap2sXmuh-XA7h5CDxZ5A
- 775 Langland-Hassan, P. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente,  
776 Eds.). Oxford University Press.
- 777 Langland-Hassan, P., Faries, F. R., Richardson, M. J., & Dietz, A. (2015). Inner speech  
778 deficits in people with aphasia. *Frontiers in Psychology*, 6, 528.
- 779 Loevenbruck, H., Grandchamp, R., Rapin, L., Nalborczyk, L., Dohen, M., Perrier, P., ...  
780 Perrone-Bertolotti, M. (2018). *Inner speech: New voices* (P. Langland-Hassan & A.  
781 Vicente, Eds.). Oxford University Press.
- 782 Lurito, J. T., Kareken, D. A., Lowe, M. J., Chen, S. H. A., & Mathews, V. P. (2000).  
783 Comparison of rhyming and word generation with fMRI. *Human Brain Mapping*, 10(3),  
784 99–106.
- 785 McGuire, P., Robertson, D., Thacker, A., David, A. S., Kitson, N., Frackowiak, R. S., &  
786 Frith, C. D. (1997). Neural correlates of thinking in sign language. *Neuroreport*, 8(3),  
787 695–698.
- 788 Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for  
789 task goals: The effects of cue type and articulatory suppression in the random task cuing  
790 paradigm. *Acta Psychologica*, 115(2-3), 123–142.
- 791 Monzel, M., Mitchell, D., Macpherson, F., Pearson, J., & Zeman, A. (2022). Aphantasia,  
792 dysikonesia, anauralia: Call for a single term for the lack of mental imagery—commentary  
793 on dance et al. (2021) and hinwar and lambert (2021). *Cortex*, 150, 149–152.  
794 https://doi.org/10.1016/j.cortex.2022.02.002
- 795 Morin, A. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente, Eds.).  
796 Oxford University Press.
- 797 Morin, A., Duhnych, C., & Racy, F. (2018). Self-reported inner speech use in university

- 798 students. *Applied Cognitive Psychology*, 32(3), 376–382.
- 799 Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames*. Retrieved from  
800 <https://CRAN.R-project.org/package=tibble>
- 801 Nash, J. C. (2014). On best practice optimization methods in R. *Journal of Statistical  
802 Software*, 60(2), 1–14. <https://doi.org/10.18637/jss.v060.i02>
- 803 Nash, J. C., & Varadhan, R. (2011). Unifying optimization algorithms to aid software  
804 system users: optimx for R. *Journal of Statistical Software*, 43(9), 1–14.  
805 <https://doi.org/10.18637/jss.v043.i09>
- 806 Nedergaard, J. S. K., Christensen, M. S., & Wallentin, M. (2021). Valence, form, and  
807 content of self-talk predict sport type and level of performance. *Consciousness and  
808 Cognition*, 89, 103102.
- 809 Nedergaard, J. S. K., Wallentin, M., & Lupyan, G. (2022). Verbal interference paradigms: A  
810 systematic review investigating the role of language in cognition. *Psychonomic Bulletin  
811 & Review*, 1–25.
- 812 Owen, W. J., Borowsky, R., & Sarty, G. E. (2004). FMRI of two measures of phonological  
813 processing in visual word recognition: Ecological validity matters. *Brain and Language*,  
814 90(1-3), 40–46.
- 815 Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal  
816 component of working memory. *Nature*, 362(6418), 342–345.
- 817 Pedersen, T. L. (2021). *Ggforce: Accelerating 'ggplot2'*. Retrieved from  
818 <https://CRAN.R-project.org/package=ggforce>
- 819 Perrone-Bertolotti, M., Rapin, L., Lachaux, J.-P., Baciu, M., & Loevenbruck, H. (2014).  
820 What is that little voice inside my head? Inner speech phenomenology, its role in  
821 cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research*,  
822 261, 220–239.
- 823 Poldrack, R. A., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., &  
824 Gabrieli, J. D. (2001). Relations between the neural bases of dynamic auditory

- 825 processing and phonological processing: Evidence from fMRI. *Journal of Cognitive*  
826 *Neuroscience*, 13(5), 687–697.
- 827 Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R.  
828 K., et al.others. (1996). Cerebral organization of component processes in reading. *Brain*,  
829 119(4), 1221–1238.
- 830 R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna,  
831 Austria: R Foundation for Statistical Computing. Retrieved from  
832 <https://www.R-project.org/>
- 833 Roebuck, H., & Lupyán, G. (2020). The internal representations questionnaire: Measuring  
834 modes of thinking. *Behavior Research Methods*, 52(5), 2053–2070.
- 835 Sarkar, D. (2008). *Lattice: Multivariate data visualization with r*. New York: Springer.  
836 Retrieved from <http://lmdvr.r-forge.r-project.org>
- 837 Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for  
838 name agreement, image agreement, familiarity, and visual complexity. *Journal of*  
839 *Experimental Psychology: Human Learning and Memory*, 6(2), 174.
- 840 Soloducha, A. (2020). What it's like living without an inner monologue. CBC News.  
841 Retrieved from <https://www.cbc.ca/news/canada/saskatchewan/inner-monologue-experience-science-1.5486969>
- 843 Terry M. Therneau, & Patricia M. Grambsch. (2000). *Modeling survival data: Extending the*  
844 *Cox model*. New York: Springer.
- 845 Thiébaut, F. I., White, S. J., Walsh, A., Klargaard, S. K., Wu, H.-C., Rees, G., & Burgess, P.  
846 W. (2015). Does faux pas detection in adult autism reflect differences in social cognition  
847 or decision-making abilities? *Journal of Autism and Developmental Disorders*, 46(1),  
848 103–112. <https://doi.org/10.1007/s10803-015-2551-1>
- 849 Wei, T., & Simko, V. (2021). *R package 'corrplot': Visualization of a correlation matrix*.  
850 Retrieved from <https://github.com/taiyun/corrplot>
- 851 Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York.

- 852        Retrieved from <https://ggplot2.tidyverse.org>
- 853    Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations.*
- 854        Retrieved from <https://CRAN.R-project.org/package=stringr>
- 855    Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors).*
- 856        Retrieved from <https://CRAN.R-project.org/package=forcats>
- 857    Wickham, H. (2021b). *Tidyr: Tidy messy data.* Retrieved from
- 858        <https://CRAN.R-project.org/package=tidyr>
- 859    Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ... Yutani,
- 860        H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.
- 861        <https://doi.org/10.21105/joss.01686>
- 862    Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation.* Retrieved from <https://CRAN.R-project.org/package=dplyr>
- 863
- 864    Wickham, H., Hester, J., & Bryan, J. (2021). *Readr: Read rectangular text data.* Retrieved
- 865        from <https://CRAN.R-project.org/package=readr>
- 866    Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'.*
- 867        Retrieved from <https://CRAN.R-project.org/package=cowplot>
- 868    Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007).
- 869        Russian blues reveal effects of language on color discrimination. *Proceedings of the*
- 870        *National Academy of Sciences*, 104(19), 7780–7785.
- 871    Winsler, A., De Leon, J. R., Wallace, B. A., Carlton, M. P., & Willson-Quayle, A. (2003).
- 872        Private speech in preschool children: Developmental stability and change, across-task
- 873        consistency, and relations with classroom behaviour. *Journal of Child Language*, 30(3),
- 874        583–608.
- 875    Xie, Y., & Allaire, J. (2022). *Tufte: Tufte's styles for r markdown documents.* Retrieved
- 876        from <https://CRAN.R-project.org/package=tufte>
- 877    Zeileis, A., & Croissant, Y. (2010). Extended model formulas in R: Multiple parts and
- 878        multiple responses. *Journal of Statistical Software*, 34(1), 1–13.

- 879 https://doi.org/10.18637/jss.v034.i01
- 880 Zeman, A. Z., Della Sala, S., Torrens, L. A., Gountouna, V. E., McGonigle, D. J., & Logie,  
881 R. H. (2010). Loss of imagery phenomenology with intact visuo-spatial task performance:  
882 A case of 'blind imagination'. *Neuropsychologia*, 48, 145–155.
- 883 Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and pipe syntax*. Retrieved  
884 from <https://CRAN.R-project.org/package=kableExtra>

**Table 4**

*Descriptive statistics of reaction time and accuracy on the task switching experiment.*

Group	Condition	Reaction time (ms)	95% CI (reaction time)	Accuracy	95% CI (Accuracy)
High-verbal	Blocked addition	2287.38	47.04	97.94	0.83
High-verbal	Color-cued switch	2774.63	61.61	95.64	1.16
High-verbal	Blocked subtraction	2527.52	53.77	97.65	0.89
High-verbal	Symbol-cued switch	2564.20	54.44	97.72	0.86
High-verbal	Un-cued switch	2678.94	59.15	94.59	1.29
Low-verbal	Blocked addition	2312.32	46.34	98.32	0.76
Low-verbal	Color-cued switch	2781.48	62.98	95.08	1.26
Low-verbal	Blocked subtraction	2572.91	55.19	97.80	0.88
Low-verbal	Symbol-cued switch	2639.81	56.00	96.72	1.03
Low-verbal	Un-cued switch	2709.74	63.84	93.19	1.47