Not everybody has an inner voice: Behavioral consequences of anendophasia

Johanne S. K. Nedergaard[1] & Gary Lupyan[2]

[1] Department of Nordic Studies and Linguistics, University of Copenhagen

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

Author Note

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

Correspondence concerning this article should be addressed to Johanne S. K. Nedergaard, Emil Holms Kanal 2, 2300 Copenhagen S, Denmark. E-mail: jskn@hum.ku.dk

Abstract

It is commonly assumed that inner speech – the experience of thought as occurring in a natural language – is both universal and ubiquitous. Recent evidence, however, suggests that similar to other phenomenal experiences like visual imagery, the experience of inner speech varies between people, ranging from constant to non-existent. We propose a name for a lack of the experience of inner speech – anendophasia – and report four studies examining some of its behavioral consequences. We found that people who report low levels of inner speech have lower performance on a verbal working memory task and have more difficulty performing rhyme judgments based on images. Task switching performance, previously linked to endogenous verbal cueing, was unaffected by differences in inner speech. Studies of anendophasia, together with aphantasia, synesthesia, and differences in autobiographical memory are providing glimpses into what may be a large space of hitherto unexplored differences in people's phenomenal experience.

Statement of Relevance

Most people say that they experience an inner voice, and that this inner voice plays an important role in their daily lives. However, a minority of people report that they do not experience such an inner voice. Until now, it has not been systematically investigated whether these self-reported differences have consequences for how people solve problems and act in the world. In this article, we found that people with less inner speech differed from people with more inner speech on some tasks that we thought would involve inner speech but not others. It is very important to understand such individual differences in inner speech use because it has consequences for how we discuss the role of inner speech generally in areas like mental health, social cognition, and self-regulation.

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

Word count: 5823

1 Introduction

Everyone, it is often said, has an inner voice, and most of our waking hours are claimed to be filled with inner speech: 'Daily, human beings are engaged in a form of inner dialogue, which enables them to high-level cognition, including self-control, self-attention and self-regulation.': (Chella & Pipitone, 2020, p. 287); 'We all hear a voice inside our brain, commonly called "inner voice", "inner speech" or referred to as "verbal thoughts" ' (Perrone-Bertolotti, Rapin, Lachaux, Baciu, and Loevenbruck (2014), p. 22). Most people do report experiencing inner speech (Alderson-Day & Fernyhough, 2015; Heavey & Hurlburt, 2008; Morin, Duhnych, & Racy, 2018) and because we often assume that our experiences mirror those of others, the majority experience comes to be viewed as universal (Lupyan, Uchiyama, Thompson, & Casasanto, 2023). The assumption that everyone has an inner voice has served as a stepping stone for research into the functions of inner speech – if everyone has it, it must be important. Speculations have ranged from the idea that natural language constitutes (at least some types of) thought (Bermúdez, 2007; Carruthers, 2002; Clark, 1998; Frankish, 2018; Gauker, 2011; Morin, 2018) to investigations of connections between inner speech and specific processes such as cognitive control (Alderson-Day & Fernyhough, 2015; Cragg & Nation, 2010; Emerson & Miyake, 2003; Morin et al., 2018). But not everyone experiences inner speech. This is attested by personal narratives such as 'What it's like living without an inner voice' (Soloducha, 2020); 'People With No Internal Monologue Explain What It's Like In Their Head' (Felton, 2020), as well as more systematic investigations both targeting variation in inner speech (Alderson-Day, Mitrenga, Wilkinson, McCarthy-Jones, & Fernyhough, 2018; Brinthaupt, 2019; Hurlburt, Heavey, & Kelsey, 2013) and auditory imagery, which has sometimes been used as a proxy for inner speech (Dawes, Keogh, Andrillon, & Pearson, 2020; Hinwar & Lambert, 2021).

1.1 The Present Study

We recruited participants differing in subjectively reported inner speech and tested them on four behavioral tasks on which performance may vary as a function of inner speech based on prior theoretical claims. The first is a rhyme judgment task: participants see pairs of images and need to indicate whether their names rhyme or not. We reasoned that although participants with low inner speech would have no trouble naming the objects, a lesser reliance on inner speech would make it harder to compare the names in memory – necessary for making a rhyme judgment (Geva, Bennett, Warburton, & Patterson, 2011; Langland-Hassan, Faries, Richardson, & Dietz, 2015). Just as visual imagery has been predicted (and sometimes found) to be linked to visual memory, we tested whether inner speech predicted memory for verbal material. We focused on memory for sets of words that were either phonologically similar and orthographically different or orthographically similar and phonologically different. Less inner speech was predicted to be associated with poorer overall memory for verbal material, but to the extent that phonological similarity creates memory confusion (Baddeley, 1966; Murray, 1968), less inner speech may be associated with a reduced phonological similarity effect. There is robust evidence that inner speech is often recruited for behavioral control when participants have to switch between different tasks (Baddeley, Chincotta, & Adlam, 2001; Emerson & Miyake, 2003; Miyake, Emerson, Padilla, & Ahn, 2004). For example, when asked to switch between adding and subtracting numbers, participants show a selective impairment if they undergo articulatory suppression, but no such impairment is found if the cues are exogenously provided (e.g., a symbol or color cue is used to inform participants whether they should add or subtract) (see Nedergaard, Wallentin, & Lupyan, 2022 for a systematic review of verbal interference effects). We reasoned that people who do not habitually use inner speech might be selectively impaired when they have to rely on self-generated cues. On the other hand, it is possible that they have learned to rely on other strategies in which case no difference would be found. Our fourth and last task involves examining category effects in perception. There is considerable evidence that language induces more categorical representations from basic perception onward (Forder & Lupyan, 2019; Perry & Lupyan, 2014; e.g., Winawer et al., 2007). In a study examining the effects of conceptual categories, Lupyan, Thompson-Schill, and Swingley (2010) showed that controlling for visual differences, people's ability to tell whether two stimuli were physically the same was affected by the categorical status of those stimuli. For example, it took longer to distinguish two cats than an equally visually similar cat and dog. We wondered whether such category effects, insofar as they may be in part induced by feedback from verbal labels, may be reduced in people with less inner speech.

2 Open Practices Statements

The experiment code and materials, data, and analysis scripts for the present study are publicly accessible at https://github.com/johannenedergaard/anendophasia. The present study was not preregistered.

3 Methods

3.1 Participants

We recruited participants online who had previously completed the Internal Representations Questionnaire (Roebuck & Lupyan, 2020) as part of unrelated studies, contacting participants with verbal factor scores < 3.5 (bottom 16%-ile) or > 4.25 (top 40%-ile) on the Verbal factor of the questionnaire which is largely centered on propensity to experience and rely on inner speech. For example, one item with a high loading on the Verbal factor was 'I think about problems in my mind in the form of a conversation with myself'. One item with a high loading on the Visual factor was 'I often enjoy the use of mental pictures to reminisce'. The percentile cut-offs were asymmetric because it was more difficult to recruit participants reporting low levels of inner speech, and because the distribution in verbal scores on the IRQ is negatively skewed. Recruiting for example the top and bottom quartiles would have resulted in a "low inner speech" group who had moderate amounts of self-stated inner speech. We received ethical approval from [redacted]. Ten participants were excluded for responding randomly, missing at least one experiment, or clearly not complying with task instructions. Our final sample included 47 participants with relatively high verbal factor scores on the IRQ and 46 participants with low verbal factor scores. The two groups were balanced in terms of age, gender, education level, dyslexia, and first language.See Table 1. Because of a technical error, demographic data is missing for one participant with less inner speech.

Table 1

Comparisons of demographic characteristics of the group with more inner speech and the group with less inner speech.

Measure	More inner speech	Less inner speech	Test for difference
Age	Median = 37; range = 18-67	Median = 39; range = 18-70	t(88.43) = -0.19; p = .849
Gender	22 female, 25 male	19 female, 26 male	$\chi^2(1) = 0.05$; p = .816
Native English-speaker	47 native speakers, 0 non-native speakers	41 native speakers, 4 non-native speakers	$\chi^2(1) = 2.49$; p = .114
Dyslexia	46 non-dyslexic, 1 self-diagnosed	44 non-dyslexic, 1 self-diagnosed	$\chi^2(1) < 0.01$; p = 1
Education level	12 high school diploma, 14 some college - no degree, 6 associate's degree, 14 bachelor's degree, 1 master's degree	1 less than high school, 14 high school diploma, 8 some college - no degree, 7 associate's degree, 11 bachelor's degree, 2 master's degree, 2 PhD, law, or medical degree	t(84.46) = -0.23; p = .815

3.2 Method: Verbal working memory

3.2.1 Materials and procedure. We used word sets from Baddeley (1966) because they were designed to be equivalent in other respects than phonological and orthographical similarity. One set contained words that were phonologically similar but not orthographically similar ("bought", "sort", "taut", "caught", and "wart"), one set contained words that were orthographically similar but not phonologically similar ("rough", "cough", "through", "dough", "bough"), and one set was a control set ("plea", "friend", "sleigh", "row", "board"). On a given trial, participants saw five words in random order from one of the sets presented sequentially in writing and were then asked to type them back in the right order. First, participants performed two practice trials with full feedback (correct/incorrect and the stimulus words – drawn from a different set than the ones used in the real experiment – shown in order). Then, participants performed 24 trials in total with eight trials from each of the three word sets. The order of both set type and words within a trial were randomized. There was no limit to how long participants could spend on reproducing the words on a given trial. See Figure 1.

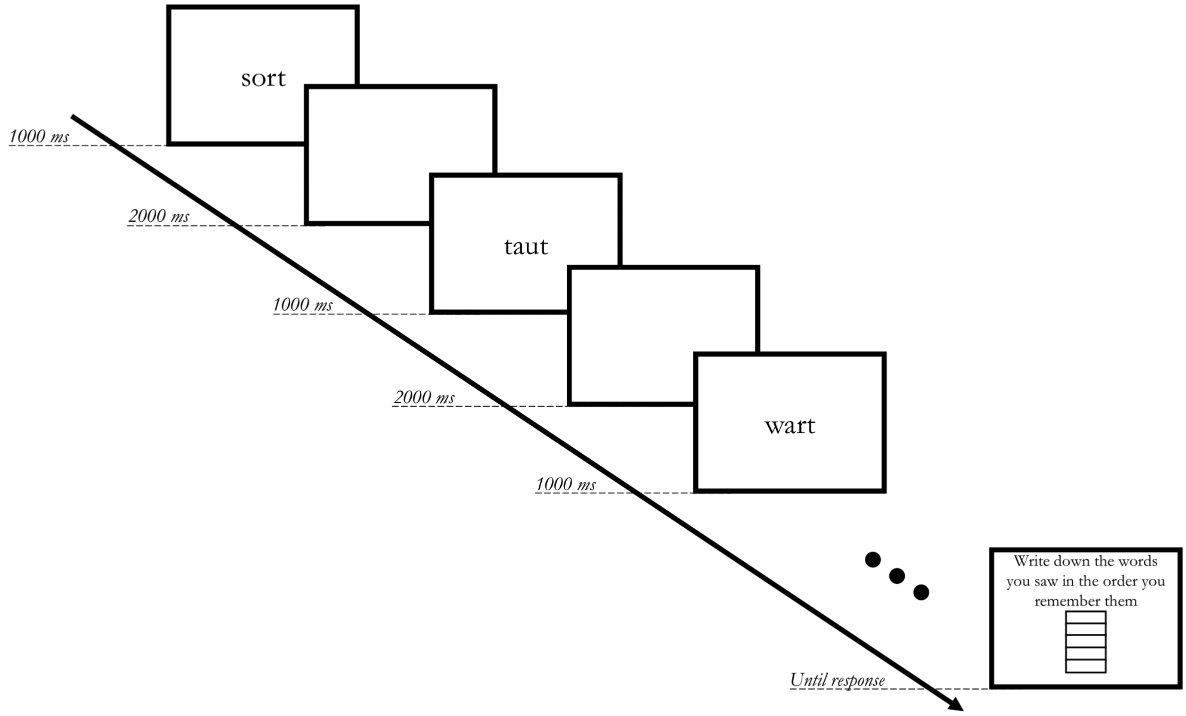

Figure 1. A sketch of the procedure in the verbal working memory experiment. In this example, the words are drawn from the phonological similarity set. Participants saw five words on each trial - three words are presented on the figure for ease of interpretation.

3.3 Method: Rhyme judgments

3.3.1 Materials and procedure. We constructed a set of rhyme pairs with 20 orthographic pairs (e.g., “sock” and “clock”) and 20 non-orthographic pairs (e.g., “drawer” and “door”). See Appendix A for the full set of images, associated words, and name agreement scores. The images were selected from the MultiPic database (Duñabeitia et al., 2018) and from Rossion and Pourtois (2004) because those image sets contained simple images (objects with no background) that had relatively high name agreement and represented the words we selected for the rhyme pairs. Participants first performed four practice trials with correct/incorrect feedback – they did not receive feedback for the remaining trials. Between each rhyme judgment trial, the screen showed a central fixation cross for either 250, 500, 750, or 1000 ms. It then showed two square black frames for 500 ms to control spatial attention – the two images then appeared simultaneously in the two squares. Participants had 5000 ms to respond to each trial and performed a total of 60 rhyme judgments in randomized order (20 orthographic rhymes, 20 non-orthographic rhymes, and 20 no-rhyme control trials). See Figure 2. Nameability scores for the images were collected from a separate set of 20 participants who were asked to label all the images. The nameability scores represent the proportion of participants who provided the target label.

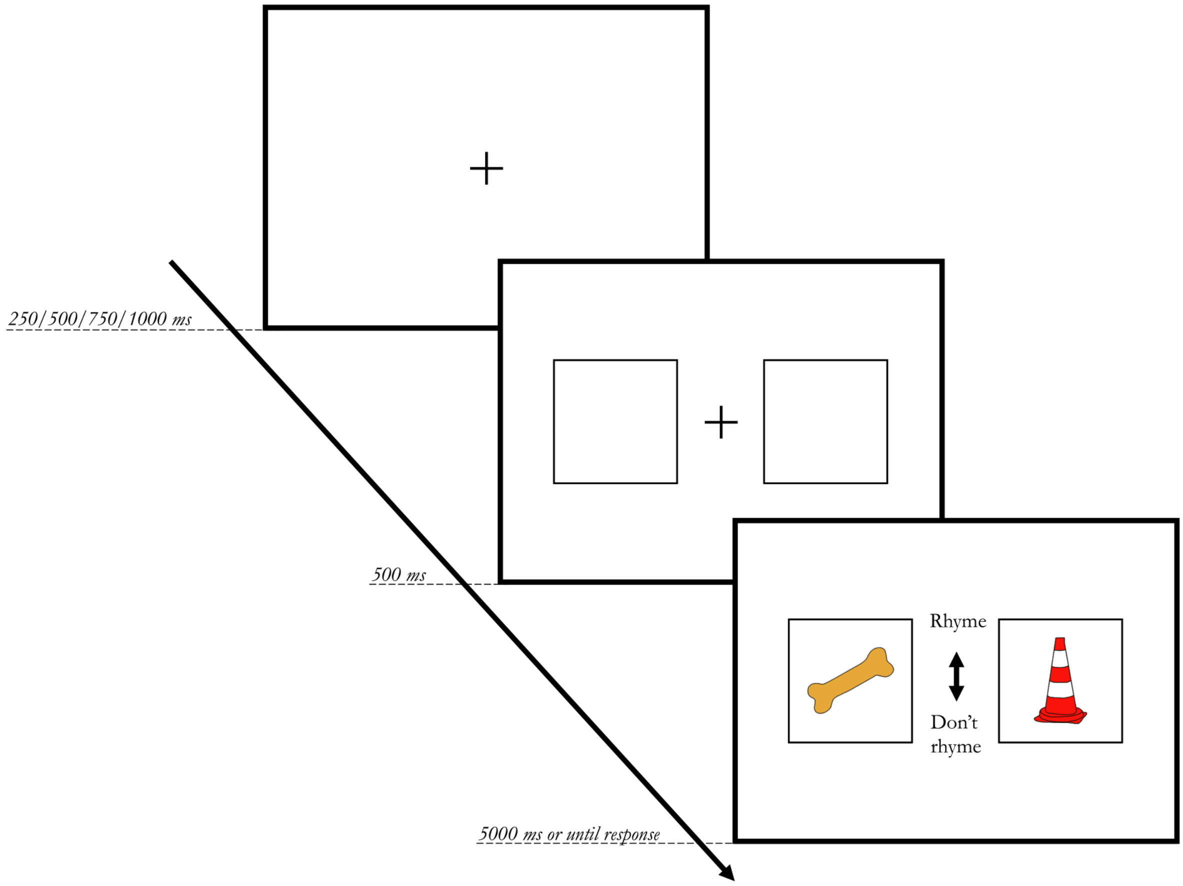

Figure 2. A sketch of a rhyme judgment trial. The stimuli here exemplify an orthographic rhyme – "bone" and "cone" – and the correct answer would therefore be "Rhyme".

3.4 Method: Task switching

3.4.1 Materials and procedure. On each block, participants were shown 30 randomly selected integers between 13 and 96 and asked to add or subtract 3 from each. All participants completed five blocks beginning with blocked addition or blocked subtraction, followed by (in a counterbalanced order) a block where problems alternated between addition and subtraction with the operation marked by color (red/blue), marked with a symbol (+/-), or not marked. The unmarked block required participants to remember which operation they had just done. For each condition, participants first solved 10 problems with correct/incorrect feedback (including feedback specific to whether the arithmetic or the operation or both were incorrect) and then 30 problems without feedback. In the switching conditions, a response counted as correct if it was the correct arithmetic and if the operation was switched from the previous trial (from addition to subtraction or vice versa). See Figure 3.

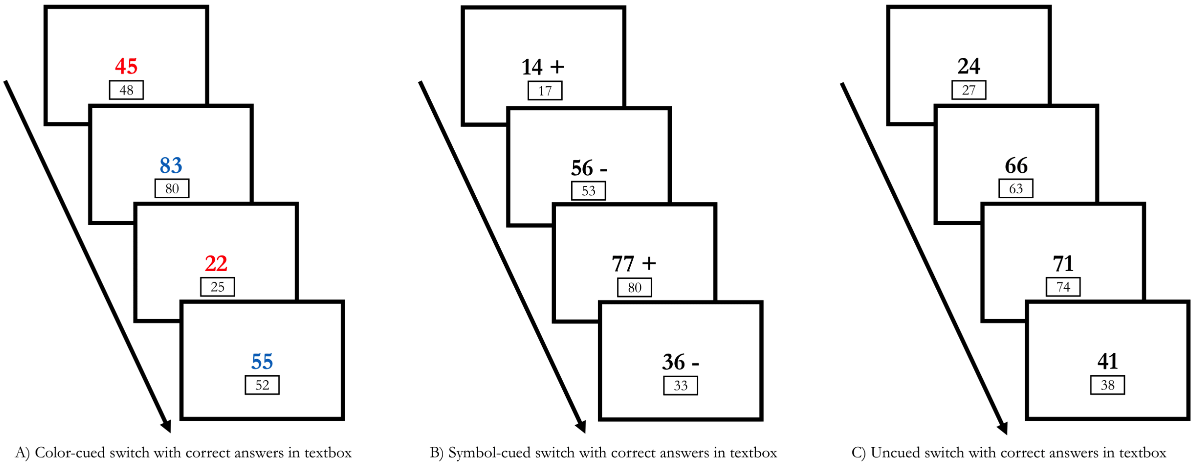

Figure 3. A sketch of the three switched conditions in the task switching experiment. Figure A shows four color-cued switch trials with correct answers, Figure B shows four symbol-cued switch trials with correct answers, and Figure C shows four un-cued switch trials with correct answers.

3.5 Method: Same/different judgments

3.5.1 Materials and procedure. This experiment used three different black silhouettes of cats and three different black silhouettes of dogs (see Figure 4).

Figure 4. The black silhouettes of cats and dogs used in the same/different judgment experiment.

There were two conditions in the experiment: a category judgment condition and an identity judgment condition. In the category judgment condition, participants were instructed to press the UP arrow key if the two animals belonged to the 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 5. 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 visual 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").

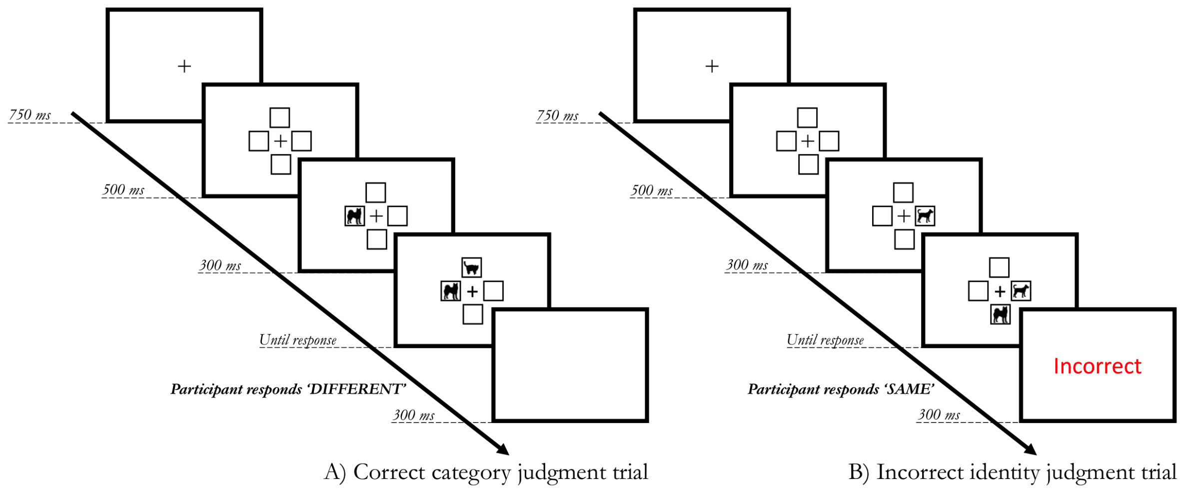

Figure 5. A sketch of the two conditions of the category judgment experiment. On Figure A, we see a correct category judgment trial where the participant responds that the cat and dog silhouettes represent different animals. On Figure B, we see an incorrect identity judgment trial where the participant responds that the two dogs are identical.

3.6 Method: Questionnaire

After completing the four experiments, participants answered custom questions about their experience with inner speech (e.g. 'How often do you have songs stuck in your head?' and '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?') and completed the Varieties of Inner Speech Questionnaire-Revised (VISQ-R) (Alderson-Day et al., 2018). See Supplemental Materials for the full set of custom questions.

3.7 Data analysis

All analyses were conducted in R version 4.1.3 (R Core Team, 2022). Participants and items (where appropriate) were modeled as random intercepts; random slopes were included for within-subject factors unless it prevented convergence. All predictors were centered. Reaction times were log-transformed to yield a more normal distribution. Accuracies were modeled using logistic regression. For ease of interpretation, the figures show the two inner speech groups as distinct but all the statistical models use verbal score (average score on the verbal representation items on the Internal Representations Questionnaire) as a continuous predictor. Error bars on all figures represent 95% confidence intervals around the mean (adjusted for repeated measures). All four experiments were conducted using custom-written software with the JavaScript package jsPsych version 6 (De Leeuw, 2015), and data and code can be found at https://github.com/johannenedergaard/anendophasia.

4 Results

4.1 Verbal working memory

4.1.1 Descriptive statistics by group: Verbal working memory.

Participants with more inner speech recalled more words correctly. This advantage was evident both when we scored only correctly ordered responses as correct as well as when we scored correctly recalled items regardless of their position (see Table 2 and Figure 6).

4.1.2 Statistical models: Verbal working memory. Participants remembered phonologically similar words significantly worse (M = 3.22) than orthographically-similar words (M = 3.62) (β = -0.72; SE = 0.08; t = -8.84; p < .001) which were in turn remembered worse than the dissimilar words (M = 3.94) (β = -0.33; SE = 0.08; t = -3.98; p < .001). Collapsing across the three types of word lists, greater inner speech was associated with better performance (β = 0.27; SE = 0.10; t = 2.60; p = .011). This effect remained significant if we disregarded the order in which participants responded, counting only whether they recalled the correct words (β = 0.19; SE = 0.08; t = 2.57; p = .012). There were no interaction effects (all p > .104), although numerically,

Table 2

Descriptive statistics by group in the verbal working memory experiment.

Group	Word set	Score (item and position)	95% CI score (item and position)	Score (position indifferent)	95% CI score (position indifferent)
More inner speech	Control set	4.19	0.13	4.51	0.08
More inner speech	Orthographic similarity set	3.72	0.14	4.18	0.10
More inner speech	Phonological similarity set	3.43	0.16	4.11	0.10
Less inner speech	Control set	3.69	0.15	4.17	0.11
Less inner speech	Orthographic similarity set	3.52	0.15	4.10	0.11
Less inner speech	Phonological similarity set	3.02	0.15	3.81	0.11

the difference was smallest for orthographically similar words (see Figure 6).

4.1.3 Strategies: Verbal working memory.

There was no difference in reported talk-out-loud strategy between the group with more inner speech (10 out of 47) and the group with less inner speech (13 out of 46) ($\chi^2(1) = 0.29$, p = .589). Nevertheless, the effect of doing so was interestingly different for the two groups as can be seen in Figure 7. The difference between the two groups' memory performance disappeared when they reported that they said the words out loud to help them remember. Participants reporting more inner speech remembered the words better, but this effect was canceled out when participants reported talking out loud to solve the task (interaction effect: β = -0.50; SE = 0.23; t = -2.19; p = .031).

4.2 Rhyme judgments

We excluded five rhyming pairs as they had below-chance performance on average for at least one group. These pairs were bin/chin, cab/crab, rake/cake, wave/cave, and park/shark. The below-chance performance was likely due to the low name agreement of at least one image in each pair (mean agreement rating for these 10 images = 0.58; range

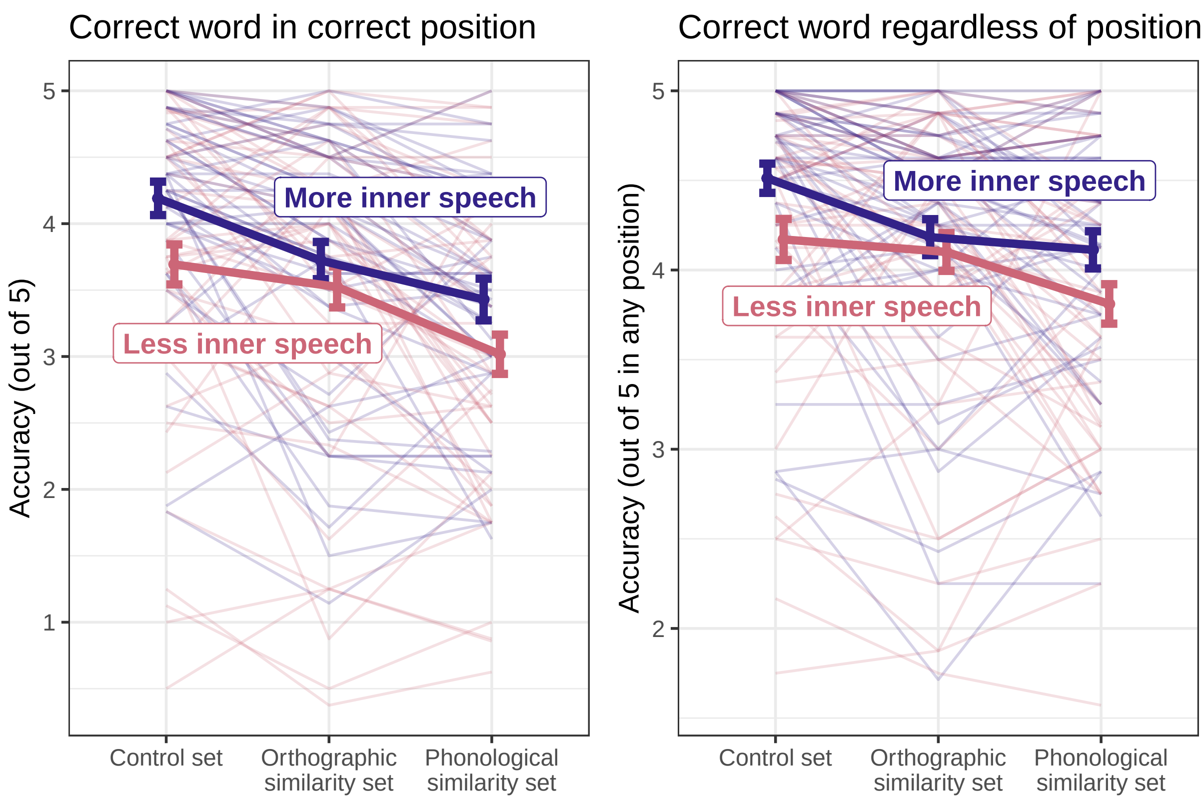

Figure 6. Score on the verbal working memory task by word set.

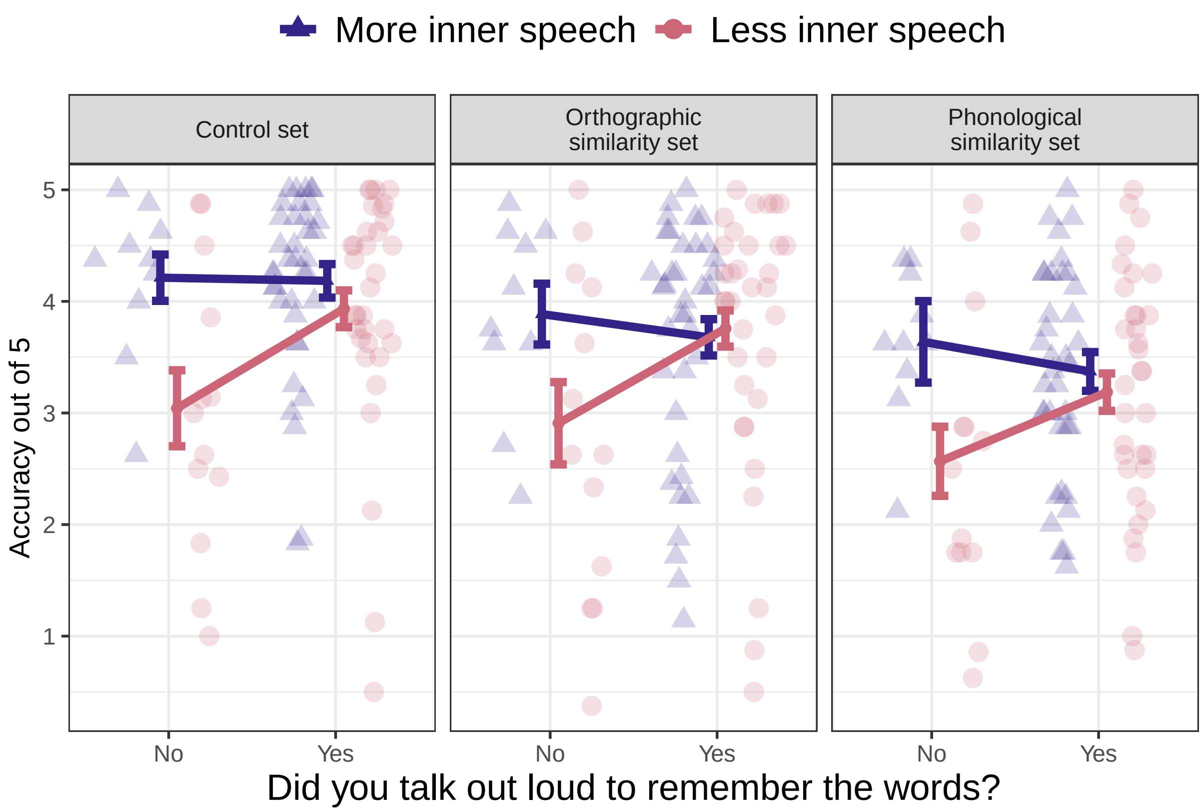

Figure 7. Verbal working memory performance by whether participants reported talking out loud to help them remember or not.

= 0.05 to 1).

4.2.1 Descriptive statistics by group: Rhyme judgments.

As can be seen in Table 3, participants with more inner speech were generally both faster and more accurate than participants with less inner speech on all three types of trials. See also Figure 8.

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)
More inner speech	Non-orthographic rhyme	1853	51	82.77	2.86
More inner speech	No rhyme	1931	53	97.52	1.36
More inner speech	Orthographic rhyme	1719	55	91.21	2.48
Less inner speech	Non-orthographic rhyme	1970	54	76.20	3.21
Less inner speech	No rhyme	2024	60	93.84	1.87
Less inner speech	Orthographic rhyme	1859	60	83.62	3.22

4.2.2 Statistical models: Rhyme judgments.

Participants took longer to make rhyme judgments on no-rhyme trials (M = 1981 ms) compared with orthographic trials (M = 1730 ms) (β = 0.12; SE = 0.04; t = 2.97; p = .005). This means that no-rhyme trials took 13% longer than orthographic trials ($e^{0.12} = 1.13$). Non-orthographic trials (M = 1821 ms) did not differ significantly from orthographic trials (β = 0.04; SE = 0.04; t = 1.11; p = .272). Trials where the presented images had higher name agreement were also faster (β = -0.04; SE = 0.02; t = -2.25; p = .029). Reported inner speech had no effect on speed of rhyme judgments (β = -0.02; SE = 0.02; t = -0.81; p = .422), and there were no interactions between rhyme type and verbal score (both $p > .298$). Verbal score and name agreement also did not interact ($p > .975$).

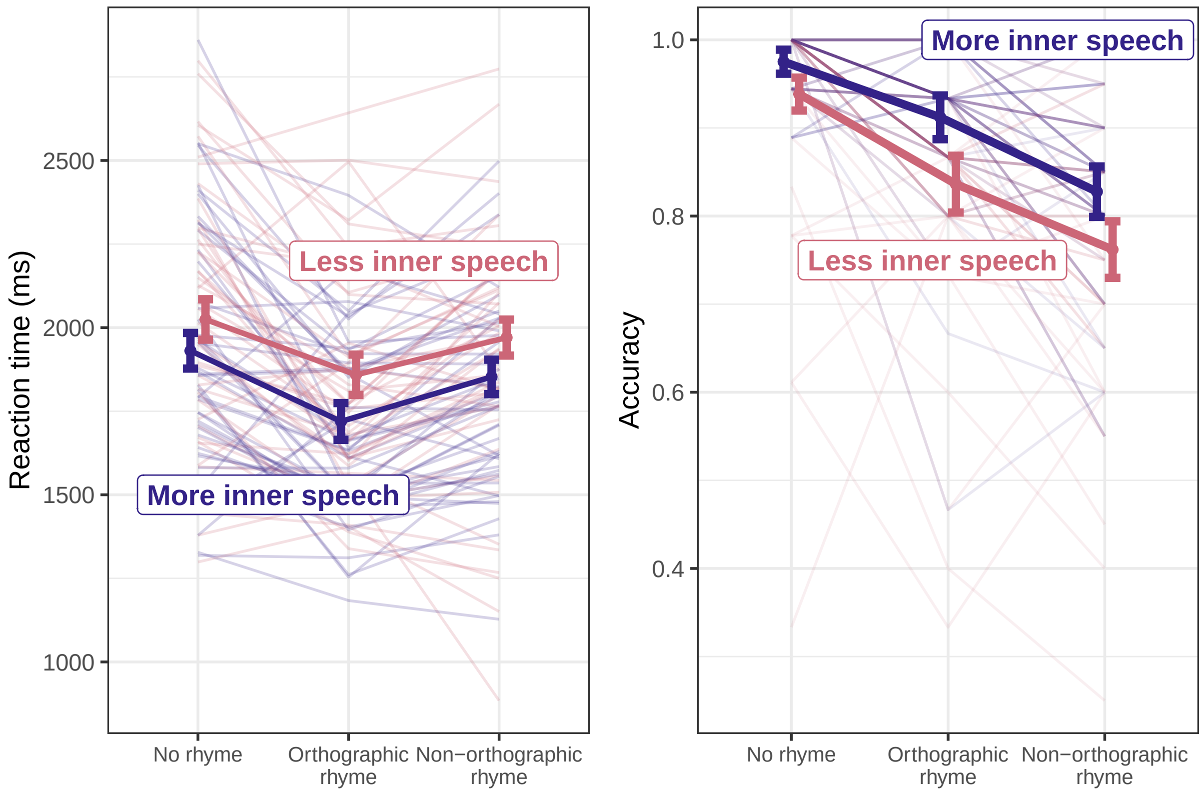

Figure 8. Reaction time and accuracy across groups by rhyme type.

Participants were more accurate on no-rhyme judgments (M = 95.7%) than on orthographic rhyme judgments (M = 87.5%) (β = 1.30; SE = 0.29; z = 4.49; $p < .001$) and less accurate on non-orthographic rhyme judgments (M = 79.5%) than on orthographic rhyme judgments (β = -0.58; SE = 0.26; z = -2.18; p = .029). A higher verbal score was associated with a higher likelihood of responding accurately (β = 0.31; SE = 0.12; z = 2.57; p = .010). Trials with images with higher name agreement were not significantly easier ($p < .139$). There was no significant interaction between rhyme type and verbal score (both $p > .311$) or between verbal score and name agreement (p = .324).

4.2.3 Strategies: Rhyme judgments. There was no significant difference between how many participants with more inner speech (23 out of 47) and how many participants with less inner speech (21 out of 46) reported that they had said the words out loud ($\chi^2(1) = 0.01$, p = .913). Nevertheless, the effect of doing so was interestingly different for the two groups as can be seen in Figure 9. Saying the words out loud diminished the accuracy advantage associated with a higher verbal score for non-orthographic rhymes (β = -0.72; SE = 0.28; z = -2.53; p = .012) and orthographic rhymes (β = -0.69; SE = 0.31; z = -2.25; p = .024) compared with no-rhyme trials. This suggests that this was the strategy that participants with more inner speech used covertly.

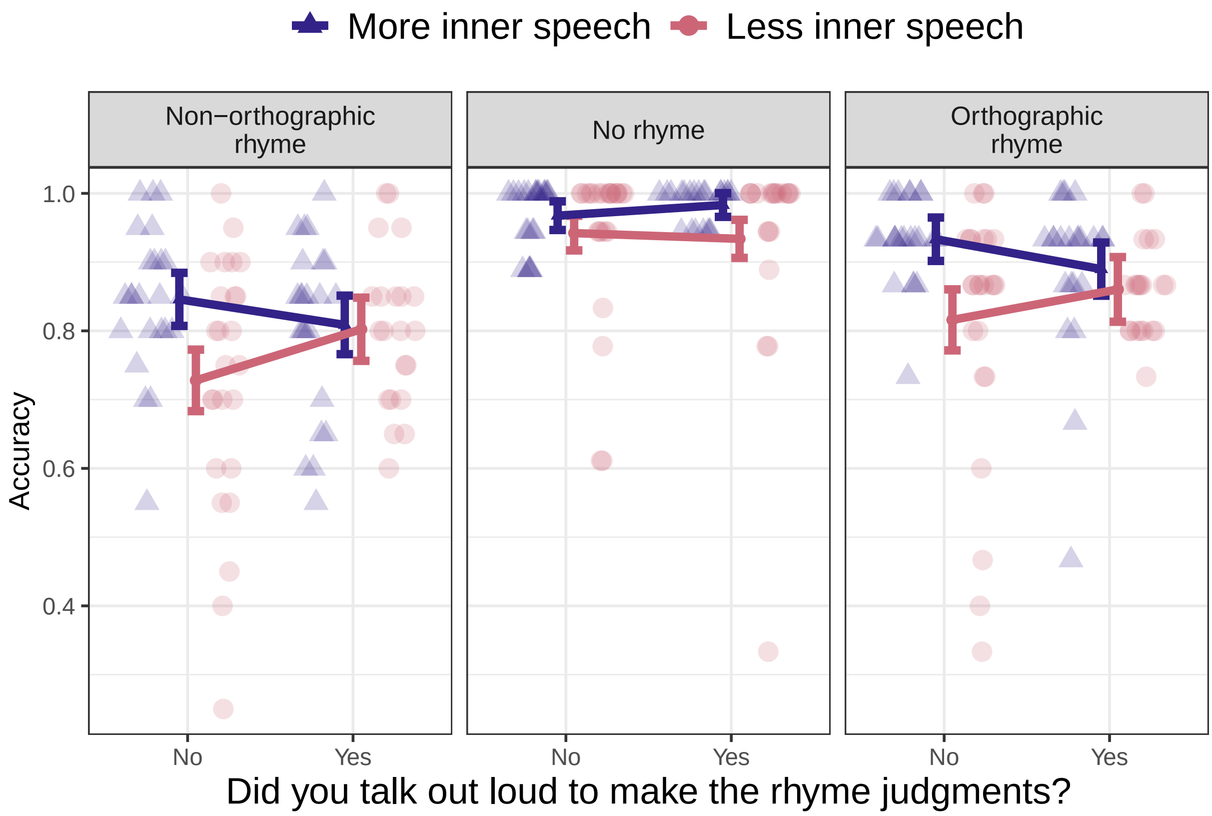

Figure 9. Reaction time and accuracy by whether participants indicated that they had talked out loud to make the rhyme judgments.

4.3 Task switching

We excluded trials over 10 seconds (0.5 % of trials). We also recalculated the accuracy measure so that any trial in the three switch conditions where participants in fact switched between adding and subtracting counted as correct (as long as the arithmetic itself was also correct). We did this to prevent a failure to switch once resulting in the remaining trials counting as incorrect.

4.3.1 Descriptive statistics: Task switching. As can be seen from Table 4 and Figure 10, accuracy was generally quite high in all conditions, and reaction times were comparable across the two groups of participants.

4.3.2 Statistical models: Task switching. Participants responded less accurately in the symbol-cued switch condition (M = 97.2%), in the color-cued switch

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)**
More inner speech	Blocked addition	2287	47	97.94	0.83
More inner speech	Color-cued switch	2775	62	95.64	1.16
More inner speech	Blocked subtraction	2528	54	97.65	0.89
More inner speech	Symbol-cued switch	2564	54	97.72	0.86
More inner speech	Un-cued switch	2679	59	94.59	1.29
Less inner speech	Blocked addition	2312	46	98.32	0.76
Less inner speech	Color-cued switch	2781	63	95.08	1.26
Less inner speech	Blocked subtraction	2573	55	97.80	0.88
Less inner speech	Symbol-cued switch	2640	56	96.72	1.03
Less inner speech	Un-cued switch	2710	64	93.19	1.47

condition (M = 95.4%), and in the un-cued switch condition (M = 93.9%) compared with the blocked addition condition (M = 98.1%) (addition versus symbol-cue: β = -0.42; SE = 0.18; z = -2.32; p = .020; addition versus color-cue: β = -0.97; SE = 0.17; z = -5.84; p < .001; addition versus un-cued: β = -1.27; SE = 0.16; z = -7.92; p < .001). Accuracy did not differ between blocked subtraction (M = 97.7%) and blocked addition (p = .239). More inner speech was not associated with different accuracy (p = .547) and there were no interaction effects between inner speech and block-type (all p > .075). Numerically, verbal score interacted with the un-cued condition and cancelled out the very slight (non-significant) reaction time advantage of a higher verbal score.

Participants responded faster in the blocked addition condition (M = 2300 ms) compared with the subtraction condition (M = 2550 ms) (β = 0.09; SE = 0.01; t = 8.41;

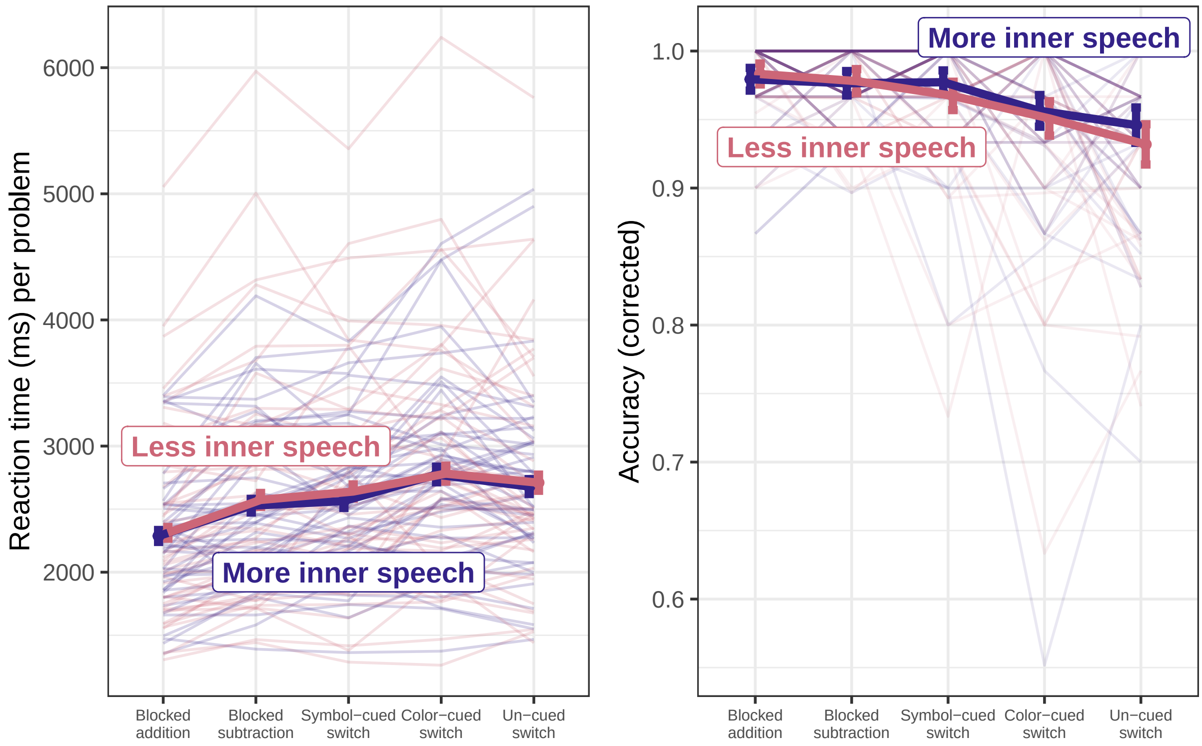

Figure 10. Reaction time and accuracy across conditions in the task switching experiment.

p < .001; regression coefficient: $e^{0.09} = 1.09$), the symbol-cued switch condition (M = 2601 ms) $\beta = 0.12$; SE = 0.01; t = 9.69; p < .001; regression coefficient: $e^{0.12} = 1.13$), the color-cued switch condition (M = 2778 ms) ($\beta = 0.19$; SE = 0.02; t = 12.23; p < .001; regression coefficient: $e^{0.19} = 1.21$), and the un-cued switch condition (M = 2694 ms) (β = 0.15; SE = 0.02; t = 9.39; p < .001; regression coefficient: $e^{0.15} = 1.16$). More reported inner speech did not predict reaction times (p = .810), and there were no interaction effects (all p > .516).

4.3.3 Strategies: Task switching. There was no significant difference between how many participants with more inner speech (20 out of 47) and how many participants with less inner speech (13 out of 46) reported that they had talked to themselves out loud during the task switching experiment ($\chi^2(1) = 1$, p = .318). There were not any obvious differences between the effects that talking out loud had on these two groups (see accuracy and reaction time Figure 11).

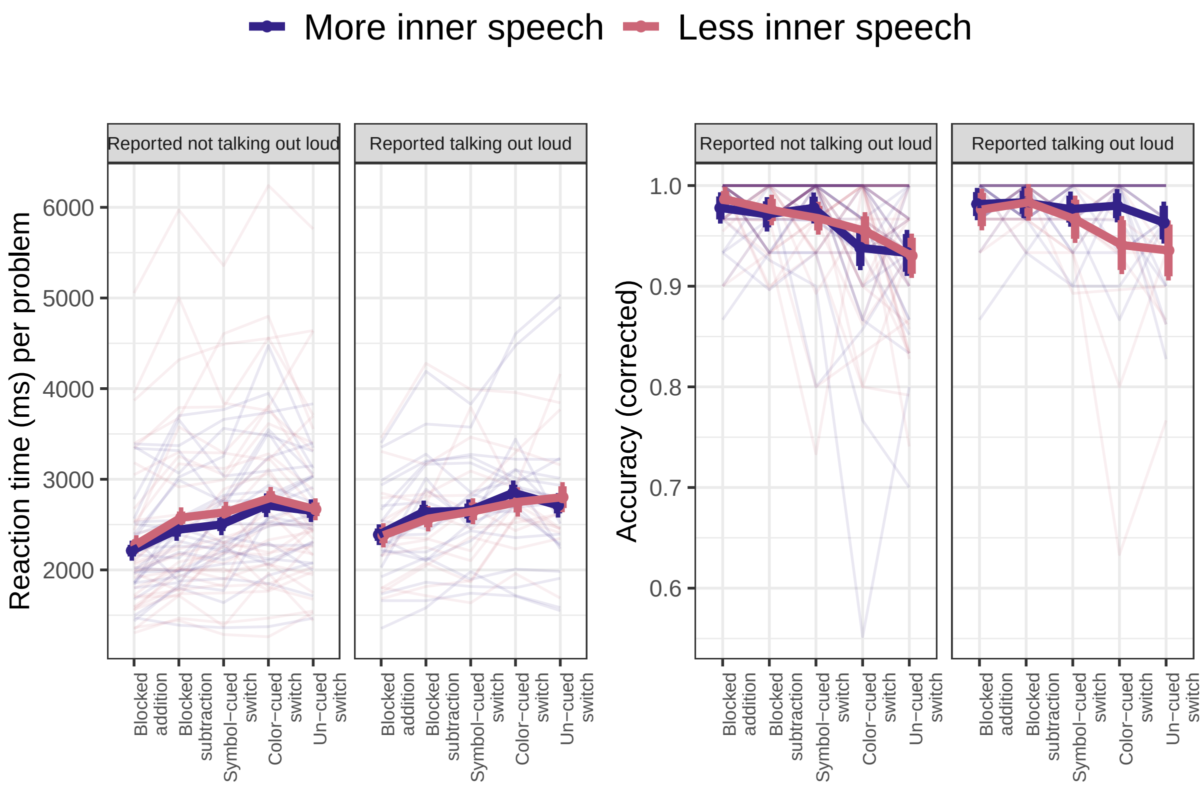

Figure 11. Reaction time (ms) and accuracy in the task switching experiment by whether participants reported talking out loud to remember the correct rule or not.

4.4 Same/different judgments

We excluded trials above 5 seconds (0.7 %) and below 200 ms (0.07 %). Generally, participants made the correct judgment on 95.53 % of trials. This did not differ between the group of participants with more inner speech (95.58 %) and the group with less inner speech (95.48 %). In subsequent analyses and plots, we only include correct trials.

4.4.1 Descriptive statistics by group: Same/different judgments. See Figure 12 for reaction times between the groups with more inner speech and less inner speech for category judgments ('do these two animals belong to the same category?') or identity judgments ('are these two animals identical?').

4.4.2 Statistical models: Same/different judgments. Identity judgments (M = 832 ms) were faster than category judgments (M = 1010 ms) (β = -0.19; SE = 0.02; t = -11.38; p < .001; regression coefficient: $e^{-0.19} = 0.83$), and a higher verbal score was not associated with faster reaction times (β = -0.03; SE = 0.02; t = -1.57; p = .120; regression coefficient: $e^{-0.03} = 0.97$).

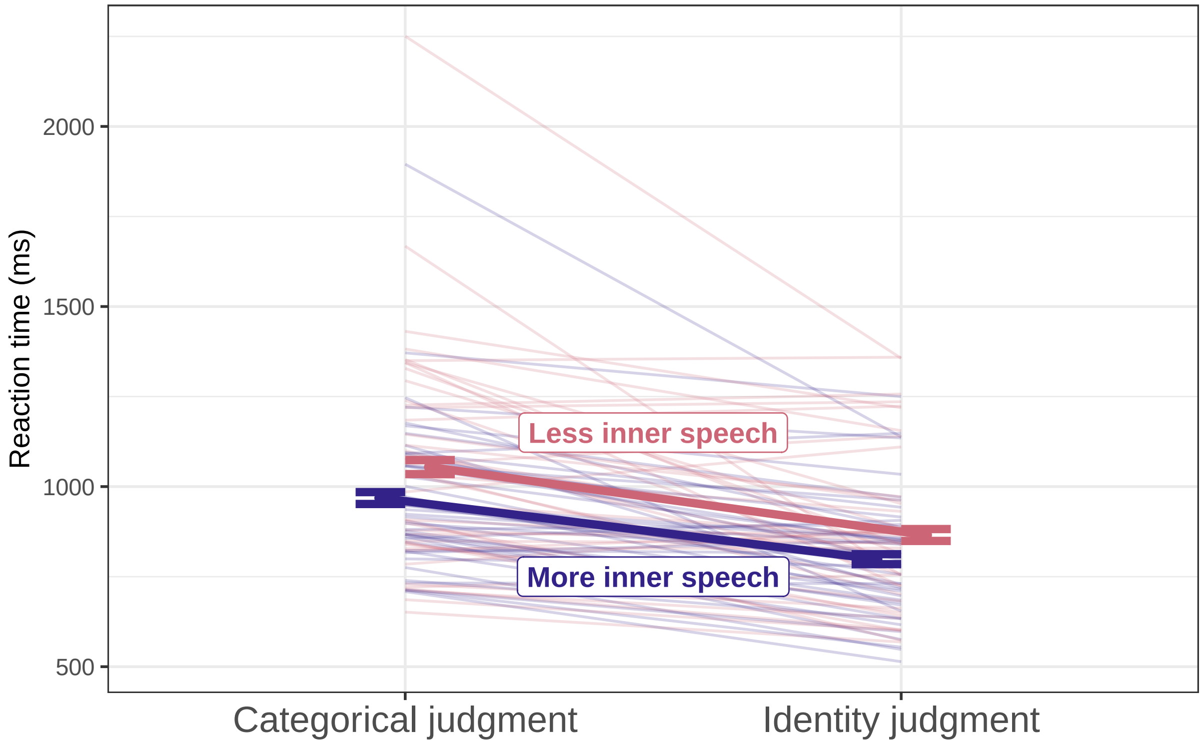

Figure 12. Reaction time in response to category or identity judgments.

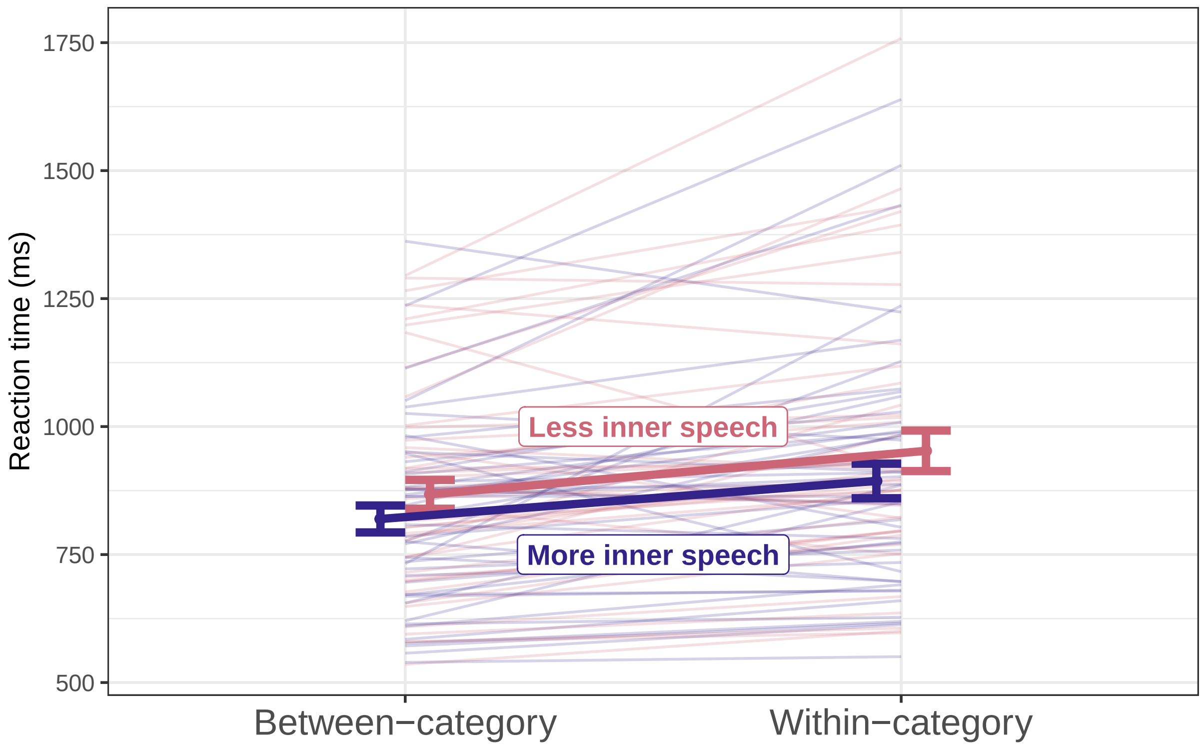

Figure 13. Reaction time on identity trials where the correct response was 'DIFFERENT' either because the two silhouettes were from different categories or different images from the same category.

The key test for this experiment was whether the two groups behaved differently when giving correct ‘DIFFERENT’ responses on identity trials when the two images belonged to the same category. That is, we expected participants with more inner speech to be slower to make correct ‘DIFFERENT’ responses when both stimuli where from the same category but physically different (i.e., dog_1 versus dog_2). See Figure 13. However, participants with more inner speech were not specifically adversely affected by the within-category interference (interaction effect: (β = 0.00; SE = 0.01; t = -0.06; p = .954). Within-category trials were generally associated with significantly slower reaction times (M = 923 ms) than between-category trials (M = 843 ms) (β = -0.08; SE = 0.01; t = -7.71; p < .001; regression coefficient: $e^{-0.08} = 0.92$). ### Strategies: Same/different judgments

There was no significant difference between how many participants with more inner speech (9 out of 47) and how many participants with less inner speech (4 out of 46) reported that they had talked to themselves out loud during the task switching experiment ($\chi^2(1) = 1.33$, p = .248). There were not any differences between the effects that talking out loud had on these two groups.

4.5 Intertask correlations

We were interested in how performance on the different tasks correlated with each other and whether these correlations were different for the two groups.

4.5.1 Intertask correlations. See Figure 14. On the plot, intertask correlations for both groups (more and less inner speech) are represented - more inner speech in the upper triangle and less inner speech in the lower triangle.

4.6 Questionnaire measures

Because of a technical error, we are missing questionnaire data from one participant from the group with less inner speech, so we here report questionnaire data from 47 participants with more inner speech and 45 participants with less inner speech. For most of our custom questions, there were notable differences in how participants from the two

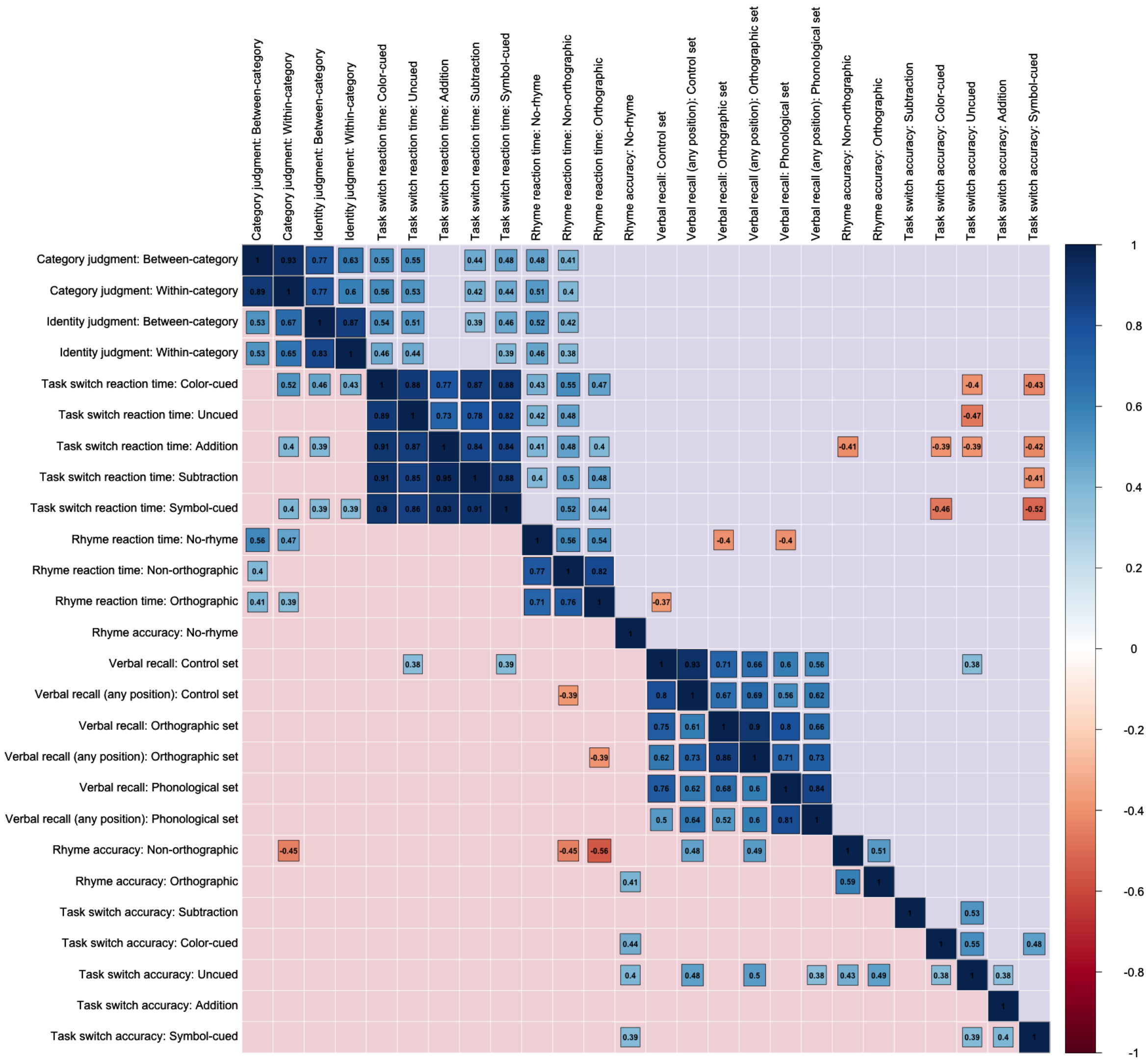

Figure 14. Intertask correlations in the total sample of participants with more and less inner speech. Colored squares represent significant correlations at p < .01. The upper triangle represents intertask correlations for the participant group with more inner speech while the lower triangle represents intertask correlations for the participant group with less inner speech.

343 groups responded. For reasons of space, however, we only report a few illustrative ones
344 here (see Appendix for plots of all the questions). The questions with the clearest
345 differences concerned rehearsing and revising conversations where the participants with
346 more inner speech reported doing so much more often than the participants with less

inner speech did (see Figure 15) (revise past conversation: $t(87.95) = 5.93$; $p < .001$; practice future conversation: $t(89.33) = 5.33$; $p < .001$). Of the VISQ factors, the IRQ verbal representation score was mostly related to the dialogicality of inner speech (see again Figure 15) ($r(90) = .70$; $p < .001$).

It was also remarkable that participants' own experience influenced how they thought other people's experience was (see Figure 16). Participants who reported more inner speech estimated that more people generally experience their thoughts in the form of a conversation with themselves ($\beta = 5.08$; SE = 2; $t = 2.55$; $p = .013$) and that more people generally hear words in their "mind's ear" when they read ($\beta = 5.09$; SE = 2.07; $t = 2.46$; $p = .016$). They did not, however, estimate that more people were able to see vivid images in their "mind's eye" ($\beta = 1.17$; SE = 2.25; $t = 0.52$; $p = .605$).

5 Discussion

Participants who report experiencing less inner speech (our sample targeted those at < 16%ile of the verbal score on the Internal Representations Questionnaire) differed in performance on several behavioral tasks. They performed worse when judging whether the names of two images rhymed, and they had poorer verbal working memory regardless of the material. Interestingly, in both the rhyming experiment and the verbal working memory experiment, performance differences between the two groups disappeared when participants reported talking out loud to solve the problems, suggesting a kind of compensatory mechanism. Inner speech differences did not predict performance in task switching which suggests that while the inner voice can be used as a behavioral self-cue, other and equally effective strategies may be available. Lastly, categorical effects on perceptual discrimination were similar for the two groups suggesting either that the categorical effects in such tasks are not language-based, or that the speeded nature of such tasks makes the use of inner speech unlikely.

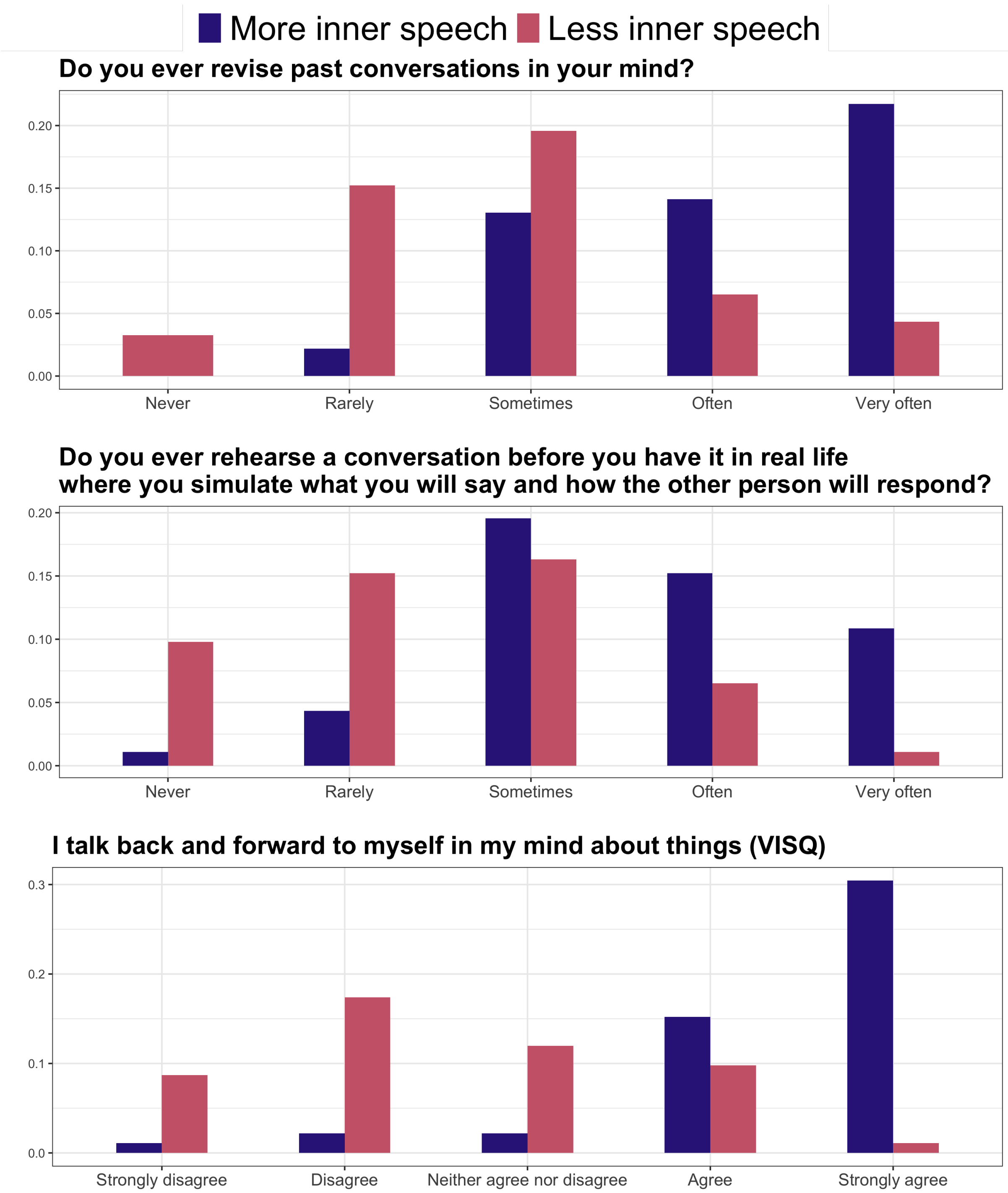

Figure 15. Grouped bar plots of proportional answers to selected custom questions concerning inner speech. Dark blue represents participants with more inner speech, and pink represents participants with less inner speech.

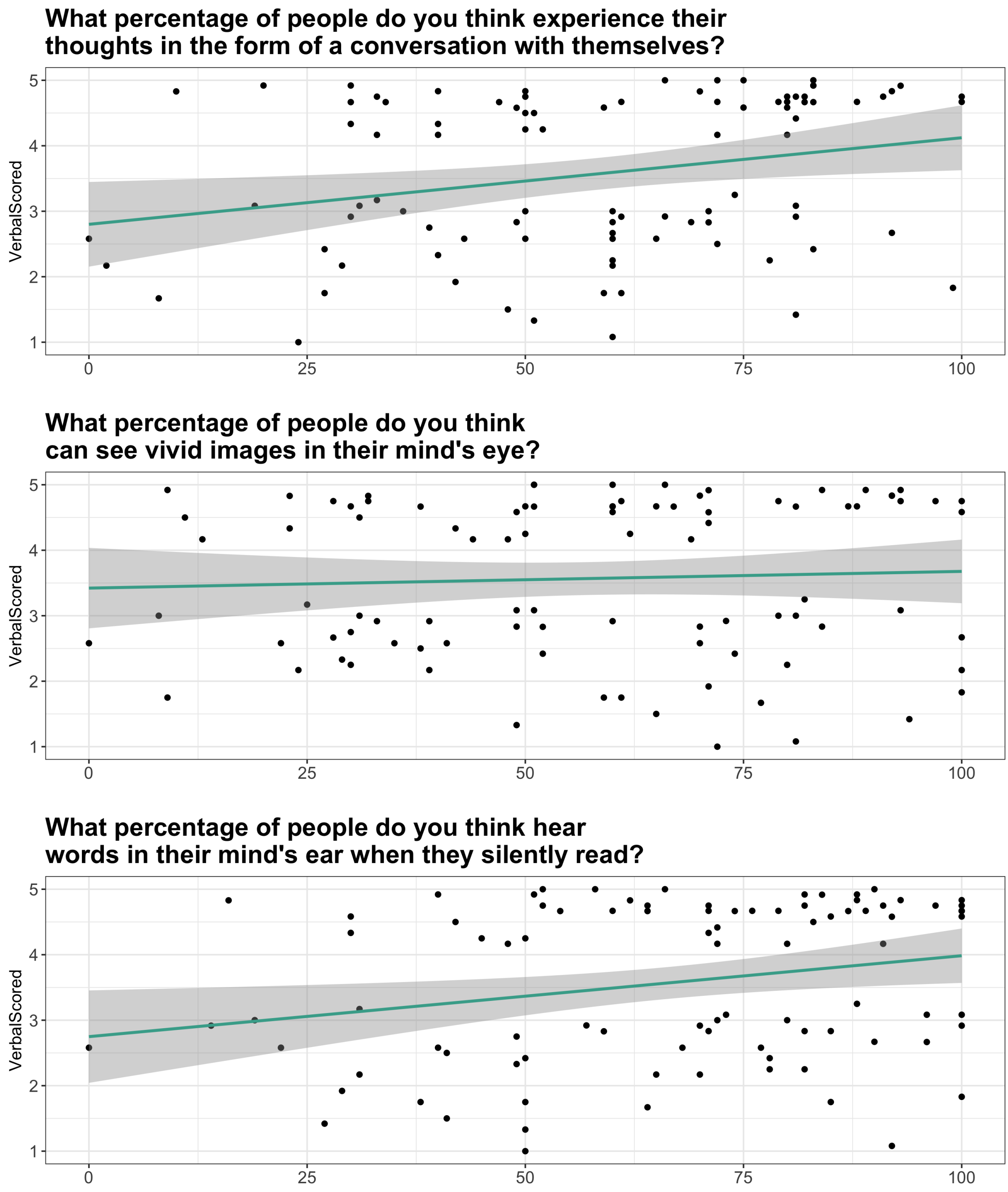

Figure 16. Scatter plots showing the correlation between verbal score on the IRQ and participants' estimates of percentages of other people with a given kind of experience.

5.1 Anendophasia: A Lack of Inner Speech

People's self-reports cannot always be taken at face value (Heavey & Hurlburt, 2008; Hurlburt, 2011; Hurlburt et al., 2013). But when people report that their

Table 5

Correlation matrix with selected variables from our custom questionnaire correlated with a dialogic item from the VISQ and Verbal and Visual scores from the IRQ.

	Simulate future conversations	Simulate past conversations	VISQ dialogic	Earworms	Others experience conversation	Others experience mind's eye	Others experience mind's ear	Verbal Score	Visual Score
Simulate future conversations	1.000								
Simulate past conversations	0.668***	1.000							
VISQ dialogic	0.548***	0.570***	1.000						
Earworms	0.498***	0.437***	0.352***	1.000					
Others experience conversation	0.409***	0.330**	0.312**	0.207*	1.000				
Others experience mind's eye	0.089	0.138	0.073	-0.052	0.403***	1.000			
Others experience mind's ear	0.266*	0.245*	0.216*	0.153	0.498***	0.452***	1.000		
Verbal Score	0.554***	0.633***	0.701***	0.461***	0.259*	0.055	0.251*	1.000	
Visual Score	0.300**	0.371***	0.208*	0.174	0.161	0.071	0.090	0.527***	1.000

experience rarely takes a verbal format, they are not just confabulating. This is evident both in the consistency of their subjective responses (Roebuck & Lupyan, 2020), and, as we report here, there are some clear behavioral correlates. When investigating unusual human experiences, it helps to have a label. For example, the coining of "aphantasia" to the lack of visual imagery (Zeman et al., 2010) is both helpful for research – providing a useful keyword – and for self-identification; its introduction led to the creation of an online community with over 50,000 members (r/aphantasia). We would therefore like to propose a name for the phenomenon of a lack of inner speech: **anendophasia**: *an* (lack) + *endo* (inner) + *phasia* (speech). This term was developed in consultation with individuals who identify as lacking inner speech and has the benefit of including the familiar Greek root *phasia* (aphasia, paraphasia, etc.). Furthermore, "endophasia" has precedent in being used to refer to inner speech (Bergounioux, 2001; Loevenbruck et al., 2018). The term also avoids subsuming inner speech under "aphantasia" (Monzel, Mitchell, Macpherson, Pearson, & Zeman, 2022) because inner speech is both auditory and articulatory in nature (whether it is better termed "inner hearing" or "inner speaking" is subject to debate) and because the linguistic properties of inner speech are not reducible to phonological properties. For these reasons, we also do not believe the previously proposed term "anauralia" is appropriate (Hinwar & Lambert, 2021).

5.2 Relations to Visual Imagery, Auditory Imagery and "Unsymbolized" Thought

Contrary to the popular belief that one is either a "verbal" or "visual" thinker (see Pashler, McDaniel, Rohrer, & Bjork, 2008 for a critical review), verbal imagery and visual imagery are in fact positively correlated (Roebuck & Lupyan, 2020). Although not the focus of the current work, our results are consistent with earlier reports of three "orientations" that all have moderate positive correlations: verbal, object/static imagery, and spatial/dynamic imagery (Blazhenkova & Kozhevnikov, 2009; Roebuck & Lupyan, 2020) suggesting a common imagery factor. Can anendophasia therefore be thought of as a lack of auditory imagery? We think not. First, many who lack inner speech report experiencing being able to hear music in their mind's ear (although they also report significantly fewer instances of "earworms"). Second, inner speech involves both auditory and articulatory-motor imagery. Second, although inner speech is often experienced as having phonological features – one of the reasons people often perceive it as speech (Langland-Hassan, 2018) – it also involves an articulatory-motor dimension (Geva, 2018; Perrone-Bertolotti et al., 2014). Paradoxically, some people also claim to experience "wordless" inner speech akin to a series of tip of the tongue states (Hurlburt et al., 2013).

When asked to reflect on what form their thoughts take, people who score low on both inner speech and visual imagery claim that they "think in concepts". What it means to "think in concepts" without relying on language is not clear. Beyond informal self-reports, the existence of such non-verbal and non-perceptual phenomenal experiences is supported by Descriptive Experience Sampling (DES) (Heavey & Hurlburt, 2008; Hurlburt & Akhter, 2006). When participants are probed at random times and asked to report on their mental states, ~22% of the time their reports are consistent with what Hurlburt has called "unsymbolized thinking". In such episodes, people feel that they think 'a particular, definite thought without awareness of that thought being conveyed as words, images, or any other symbols' (Heavey & Hurlburt, 2008, p. 802). Unsymbolized thinking is a slippery construct that tends to be defined in terms of what it is not. For example, Hurlburt and Akhter (2008) say that it is experienced as being 'a thinking, not a feeling, not an intention, not an intimation, not a kinesthetic event, not a bodily event' (p. 1366). A telling example is a participant wondering if her friend will arrive in a car or pickup truck, but not experiencing any words or images. The question is a single undifferentiated whole. It is possible that unsymbolized thinking is subserved by the same verbal and perceptual processes, but with weak or absent conscious imagery (Vicente & Martinez-Manrique, 2016). Alternatively, it may correspond to a genuinely different form of experience in which people entertain more abstract conceptual representations which are less accessible to people with higher levels of inner speech and imagery.

5.3 Limitations

One limitation of our work is its reliance on wholly subjective questions for measuring inner speech. Considering that our focus is on differences in phenomenology, this is appropriate. At the same time, there is reason to be skeptical of people's assessments of their inner experiences. People can be wrong about what they think they experience (Hurlburt & Schwitzgebel, 2011). It would be therefore helpful to supplement subjective assessments with objective ones of the sort becoming possible for differences in visual imagery (Kay, Keogh, Andrillon, & Pearson, 2022). Another limitation is the remaining possibility that differences we ascribe to inner speech come from something else such as differences in conscientiousness. We believe this is unlikely since we saw examples of specific conditions where there were no differences between the two groups (e.g., no-rhyme pairs, orthographically similar words, and all conditions in the task switching experiment). Given the wide age range and relatively balanced gender and education distributions, we believe our results to be generalizable. It is, however, worth noting that the vast majority of our participants were native English speakers and that the results could conceivably be different for other languages. Lastly, while the term "anendophasia" connotes *lack* of inner speech, many of the participants in our "low inner speech" group reported having *some* inner speech. Screening a larger group to identify people who do not endorse having *any* inner speech would help us see if the cognitive consequences of having less inner speech are continuous with having none.

6 Conclusion

Not everyone experiences inner speech. We proposed a name for a lack of inner speech: anendophasia. People who experience less inner speech were worse at making rhyme judgments in response to images and remembering a list of words. Task switching performance was not, however, either slower or less accurate. Taken together, our experiments suggest that there are real behavioral consequences of experiencing less or more inner speech, and that these differences may often be masked because people with anendophasia use alternative strategies.

7 References

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

Alderson-Day, B., Mitrenga, K., Wilkinson, S., McCarthy-Jones, S., & Fernyhough, C. (2018). The varieties of inner speech questionnaire–revised (VISQ-r): Replicating and refining links between inner speech and psychopathology. *Consciousness and Cognition*, *65*, 48–58.

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

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology: General*, *130*(4), 641.

Bergounioux, G. (2001). Endophasie et linguistique [décomptes, quotes et squelette]. *Langue Francaise*, *132*, 106–124.

Bermúdez, J. L. (2007). *Thinking without words*. Oxford University Press.

Blazhenkova, O., & Kozhevnikov, M. (2009). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *23*(5), 638–663.

Brinthaupt, T. M. (2019). Individual differences in self-talk frequency: Social isolation and cognitive disruption. *Frontiers in Psychology*, *10*, 1088.

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences*, *25*(6), 657–674.

Chella, A., & Pipitone, A. (2020). A cognitive architecture for inner speech. *Cognitive Systems Research*, *59*, 287–292.

Clark, A. (1998). *Language and thought: Interdisciplinary themes* (P. Carruthers & J. Boucher, Eds.). Cambridge University Press.

Cragg, L., & Nation, K. (2010). Language and the development of cognitive control. *Topics in Cognitive Science*, *2*(4), 631–642.

Dawes, A. J., Keogh, R., Andrillon, T., & Pearson, J. (2020). A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, *10*(1), 1–10.

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six european languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, *48*(1), 148–168.

Felton, J. (2020). People with no internal monologue explain what it's like in their head. IFLScience. Retrieved from https://www.iflscience.com/people-with-no-internal-monologue-explain-what-its-like-in-their-head-57739

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

Frankish, K. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente, Eds.). Oxford University Press.

Gauker, C. (2011). *Words and images: An essay on the origin of ideas*. Oxford University Press, Oxford.

Geva, S. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente, Eds.). Oxford University Press.

Geva, S., Bennett, S., Warburton, E. A., & Patterson, K. (2011). Discrepancy between inner and overt speech: Implications for post-stroke aphasia and normal language processing. *Aphasiology*, *25*(3), 323–343.

Heavey, C. L., & Hurlburt, R. T. (2008). The phenomena of inner experience. *Consciousness and Cognition*, *17*(3), 798–810.

Hinwar, R. P., & Lambert, A. J. (2021). Anauralia: The silent mind and its association with aphantasia. *Frontiers in Psychology*, *12*.

https://doi.org/10.3389/fpsyg.2021.744213

Hurlburt, R. T. (2011). *Investigating pristine inner experience.* Cambridge University Press. https://doi.org/10.1017/cbo9780511842627

Hurlburt, R. T., & Akhter, S. A. (2006). The descriptive experience sampling method. *Phenomenology and the Cognitive Sciences*, *5*(3), 271–301.

Hurlburt, R. T., & Akhter, S. A. (2008). Unsymbolized thinking. *Consciousness and Cognition*, *17*(4), 1364–1374.

Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, *22*(4), 1477–1494. https://doi.org/10.1016/j.concog.2013.10.003

Hurlburt, R. T., & Schwitzgebel, E. (2011). *Describing inner experience?: Proponent meets skeptic.* Mit Press.

Kay, L., Keogh, R., Andrillon, T., & Pearson, J. (2022). The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength. *Elife*, *11*, e72484.

Langland-Hassan, P. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente, Eds.). Oxford University Press.

Langland-Hassan, P., Faries, F. R., Richardson, M. J., & Dietz, A. (2015). Inner speech deficits in people with aphasia. *Frontiers in Psychology*, *6*, 528.

Loevenbruck, H., Grandchamp, R., Rapin, L., Nalborczyk, L., Dohen, M., Perrier, P., … Perrone-Bertolotti, M. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente, Eds.). Oxford University Press.

Lupyan, G., Thompson-Schill, S. L., & Swingley, D. (2010). Conceptual penetration of visual processing. *Psychological Science*, *21*(5), 682–691.

Lupyan, G., Uchiyama, R., Thompson, B., & Casasanto, D. (2023). Hidden differences in phenomenal experience. *Cognitive Science*, *47*(1), e13239.

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

Monzel, M., Mitchell, D., Macpherson, F., Pearson, J., & Zeman, A. (2022). Aphantasia, dysikonesia, anauralia: Call for a single term for the lack of mental imagery–commentary on dance et al. (2021) and hinwar and lambert (2021). *Cortex*, *150*, 149–152. https://doi.org/10.1016/j.cortex.2022.02.002

Morin, A. (2018). *Inner speech: New voices* (P. Langland-Hassan & A. Vicente, Eds.). Oxford University Press.

Morin, A., Duhnych, C., & Racy, F. (2018). Self-reported inner speech use in university students. *Applied Cognitive Psychology*, *32*(3), 376–382.

Murray, D. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, *78*(4p1), 679–684.

Nedergaard, J. S. K., Wallentin, M., & Lupyan, G. (2022). Verbal interference paradigms: A systematic review investigating the role of language in cognition. *Psychonomic Bulletin & Review*, 1–25.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, *9*(3), 105–119.

Perrone-Bertolotti, M., Rapin, L., Lachaux, J.-P., Baciu, M., & Loevenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research*, *261*, 220–239.

Perry, L. K., & Lupyan, G. (2014). The role of language in multi-dimensional categorization: Evidence from transcranial direct current stimulation and exposure to verbal labels. *Brain and Language*, *135*, 66–72.

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Roebuck, H., & Lupyan, G. (2020). The internal representations questionnaire: Measuring modes of thinking. *Behavior Research Methods*, *52*(5), 2053–2070.

Rossion, B., & Pourtois, G. (2004). Revisiting snodgrass and vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217–236.

Soloducha, A. (2020). What it's like living without an inner monologue. CBC News. Retrieved from https://www.cbc.ca/news/canada/saskatchewan/inner-monologue-experience-science-1.5486969

Vicente, A., & Martinez-Manrique, F. (2016). The nature of unsymbolized thinking. *Philosophical Explorations*, *19*(2), 173–187.

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

Zeman, A. Z., Della Sala, S., Torrens, L. A., Gountouna, V. E., McGonigle, D. J., & Logie, R. H. (2010). Loss of imagery phenomenology with intact visuo-spatial task performance: A case of 'blind imagination'. *Neuropsychologia*, *48*, 145–155.