Not everybody has an inner voice: Behavioral consequences of anendophasia

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

[1] Department of Linguistics, Cognitive Science and Semiotics, Aarhus University

[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, Jens Chr. Skous Vej 2, 8000 Aarhus C, Denmark. E-mail: nedergaard@cc.au.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 experiments 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, as was a visual discrimination task. We also report results of a questionnaire showing anendophasia to be associated with a range of experiential differences ranging from experiencing earworms to memory for conversations. We discuss our findings in relation to aphantasia, condensed versus expanded inner speech, and unsymbolized thinking.

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

Word count: 7167

1 Introduction

It is frequently claimed that everyone has an inner voice, and that most of our waking hours are filled with internal monologue [‘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 and 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. 221]. Recent evidence – both anecdotal accounts and more systematic investigations – challenge this view. In mass media, the topic has received much attention in viral Twitter threads (e.g., @KylePlantEmoji, 2020, see Figure 1) as well as in articles such as ‘What it’s like living without an inner voice’ (Soloducha, 2020) and ‘People With No Internal Monologue Explain What It’s Like In Their Head’ (Felton, 2020). Systematic investigations have focused on auditory imagery as a proxy for inner speech (Dawes, Keogh, Andrillon, & Pearson, 2020; Hinwar & Lambert, 2021) and found that auditory imagery, like visual imagery, varies from entirely absent to ubiquitous across individuals.

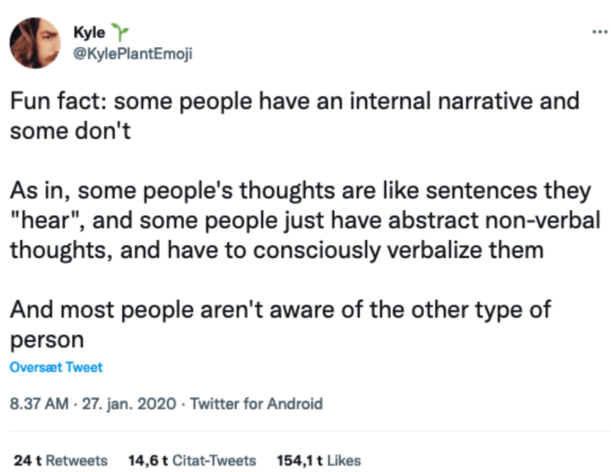

Figure 1. Viral tweet from @KylePlantEmoji about the presence or absence of inner speech. Screenshot from November 17th 2022.

Judging by these accounts, there are important differences in the extent to which people experience an inner voice. Whether these differences in experience result in differences in behavior is still an open question. We explore this intriguing possibility in the present study. If there *are* differences, this helps us understand the extent to which people's cognition may be differentially guided by language. If there are *no* differences, this could either mean that the measure of inner speech is invalid (i.e., people reporting more or less inner speech in fact have similar experiences but interpret questions concerning inner speech differently), or that differences in inner speech have no bearing on the behavioral measures in question, or that people who experience no inner speech do not differ in accuracy or speed because they rely on different processes or strategies than inner speech.

The assumption that everyone has an inner voice has served as a stepping stone for much research into the functions of inner speech – if everyone has it, it must be important. This importance includes claims that inner speech constitutes (at least some types of) thought (Bermúdez, 2007; Carruthers, 2002; Clark, 1998; Frankish, 2018; Gauker, 2011; Morin, 2018) and theories that inner speech is crucially involved in behavioral control (Alderson-Day & Fernyhough, 2015; Cragg & Nation, 2010; Emerson & Miyake, 2003; Morin, Duhnych, & Racy, 2018).

1.1 Parallels with condensed inner speech and unsymbolized thinking

What do people mean when they say they do not experience inner speech? Anecdotally, some report on internet fora that their thinking takes place largely in the visuospatial modality while another common description is 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 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 their inner experience, their reports are often consistent with what Hurlburt and colleagues have called "unsymbolized thinking" (around 22 % of experience prompts). In such episodes, people feel that they think ‘a particular, definite thought without the awareness of that thought’s being conveyed words, images, or any other symbols’ (Heavey & Hurlburt, 2008, p. 802). Unsymbolized thinking is a slippery phenomenon mostly characterized with negative definitions. 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. Instead, the question is a single, undifferentiated whole. It is possible that unsymbolized thinking is continuous with inner speech with weak or absent conscious imagery since it to some extent appears to have similar semantic and syntactic structures as language (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.

Descriptive Experience Sampling has yielded another finding that is potentially relevant for what is experienced as a lack of inner speech: “wordless” inner speech (Hurlburt, Heavey, & Kelsey, 2013), akin to a series of tip of the tongue states. In such episodes, people often report experiencing the pace, rhythm, and linear sequence of speech without the experience of hearing or speaking words. The idea that inner speech may vary in how closely tied it is to audition and articulation plays an important role in several different conceptualizations of inner speech (Fernyhough, 2004; Grandchamp et al., 2019; Oppenheim & Dell, 2010). For example, the developmental psychologist Vygotsky thought that adult inner speech is an internalized form of children’s overt speech, and that inner speech during this internalization is transformed to be more condensed in terms of both form and meaning. Vygotsky thought that the most condensed form of inner speech could be thought of as ‘thinking in pure meanings’ and thus be abstracted away from both phonological and articulatory specification (Vygotsky, 1962). Some theorists have suggested that the degree to which inner speech is experienced as condensed or expanded varies across both individuals and situations (Fernyhough, 2004; Grandchamp et al., 2019), for example under higher cognitive demands or social isolation (Brinthaupt, 2019).

1.2 Parallels with aphantasia

That there are differences in subjective reports of inner experience is not a new finding, nor is the idea that such differences may result in subtle behavioral changes. In recent years, a very similar phenomenon to internal verbal experience has gained much attention, namely the presence or absence of visual imagery. In a 2010 article, Zeman and colleagues termed the inability to engage in visual imagery “aphantasia” and reported that two thirds of the participants with aphantasia had difficulties with autobiographical memory (Zeman et al., 2010). Generally, participants with aphantasia report weak or non-existing ability to visualize “in the mind’s eye” (Dawes et al., 2020; Keogh & Pearson, 2018) and may display poorer visual working memory performance than control participants (Jacobs, Schwarzkopf, & Silvanto, 2018) although this is not always the case (Keogh, Wicken, & Pearson, 2021). The conflicting findings about consequences of aphantasia in terms of working memory abilities have prompted a discussion of whether aphantasia represents a metacognitive deficit rather than difficulties with mental visual imagery. However, recent findings suggest that a more likely explanation is that people with aphantasia simply use different strategies to solve tasks that would normally require visual imagery. For example, Keogh, Wicken, and Pearson (2021) found that participants with aphantasia performed at the same level as control participants on visual working memory tasks. There were, however, marked differences in the reported strategies used by participants with aphantasia who reported rehearsing patterns verbally or ‘using ideas and semantics’ to remember the test items. Additionally, performance levels on a number working memory task and a visual working memory task were correlated for participants with aphantasia but not for control participants. This suggests that control participants used different strategies for the two types of tasks (one is traditionally thought to occupy verbal resources while the other is thought to use visual working memory resources) while participants with aphantasia may have used similar strategies for the two different tasks. The finding that differences in strategies are likely to mask differences in visualizing ability is important for research in inner speech as well. We might see comparable performance levels due to compensatory strategies that would then mask differences in mental verbalizing abilities.

1.3 The present study

Taking inspiration from aphantasia research on visual working memory, we can also test the **verbal working memory** performance of people reporting little to no inner speech. This allows us to both test whether verbal working memory and reported inner speech use are related and explore the possible compensatory strategies. In particular, we might expect difficulties with verbal working memory tasks requiring a high degree of phonological precision (Jacobs et al., 2018). In the present study, 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. To the extent that phonological similarity makes recall more difficult (Baddeley, 1966; Murray, 1968), less inner speech may be associated with a reduced phonological similarity effect.

To further probe participants' internal verbal representations, we use a **rhyme judgment** task (Geva, Bennett, Warburton, & Patterson, 2011; Langland-Hassan, Faries, Richardson, & Dietz, 2015) where participants see two images and have to judge whether the associated words rhyme or not. Presumably, this would require them to internally verbalize. Importantly, we need to include both orthographic rhymes (such as "boat" and "moat") and non-orthographic rhymes (such as "sleigh" and "hay") as participants could otherwise make rhyme judgments by visualizing the orthographic representations of the words. We reasoned that although participants reporting low levels of inner speech would have no trouble naming the objects, less reliance on inner speech would make it harder to compare the names in memory – necessary for making a rhyme judgment.

There is robust evidence that inner speech is often recruited for behavioral control in **task switching** paradigms where participants have to switch between different task rules (Baddeley, Chincotta, & Adlam, 2001; Emerson & Miyake, 2003; Goschke, 2000; 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 the verbal interference literature). 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.

There is considerable evidence that language induces more **categorical representations** from basic perception onward (e.g. Forder & Lupyan, 2019; Perry & Lupyan, 2014; 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 (Lupyan, 2012), may be reduced in people with less inner speech.

Thus, in the present study, we explore individual differences related to reported inner speech in four behavioral tasks: verbal working memory, rhyme judgment, task switching, and categorical and perceptual visual discrimination.

2 Methods

2.1 Participants

We recruited participants 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. The percentiles were asymmetrical 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 instead would have resulted in a "low inner speech" group who did not in fact have very low verbal representation scores. We received ethical approval from the University of Wisconsin-Madison. 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.

2.2 Method: Verbal working memory

2.2.1 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

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

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 2.

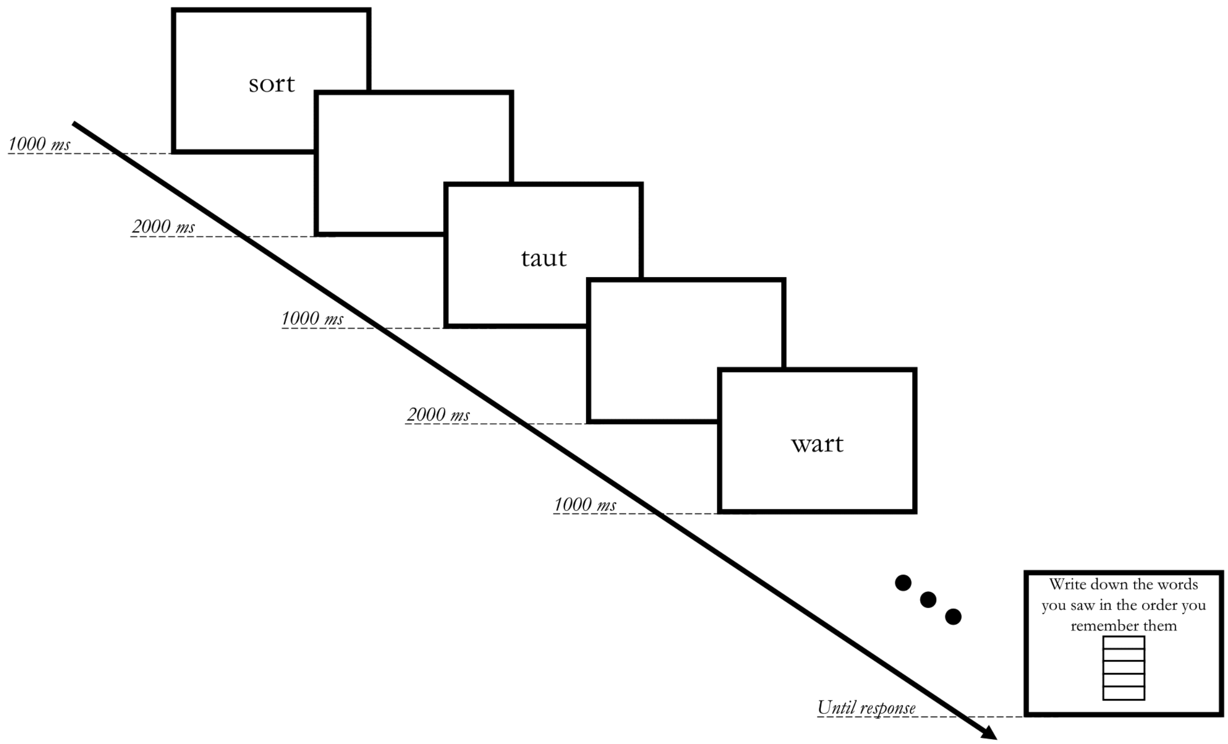

Figure 2. A sketch of the procedure in the verbal working memory experiment. In this example, the words are drawn from the phonological similarity set. Participants saw five words on each trial - three words are presented on the figure for ease of interpretation.

2.3 Method: Rhyme judgments

2.3.1 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 3. 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.

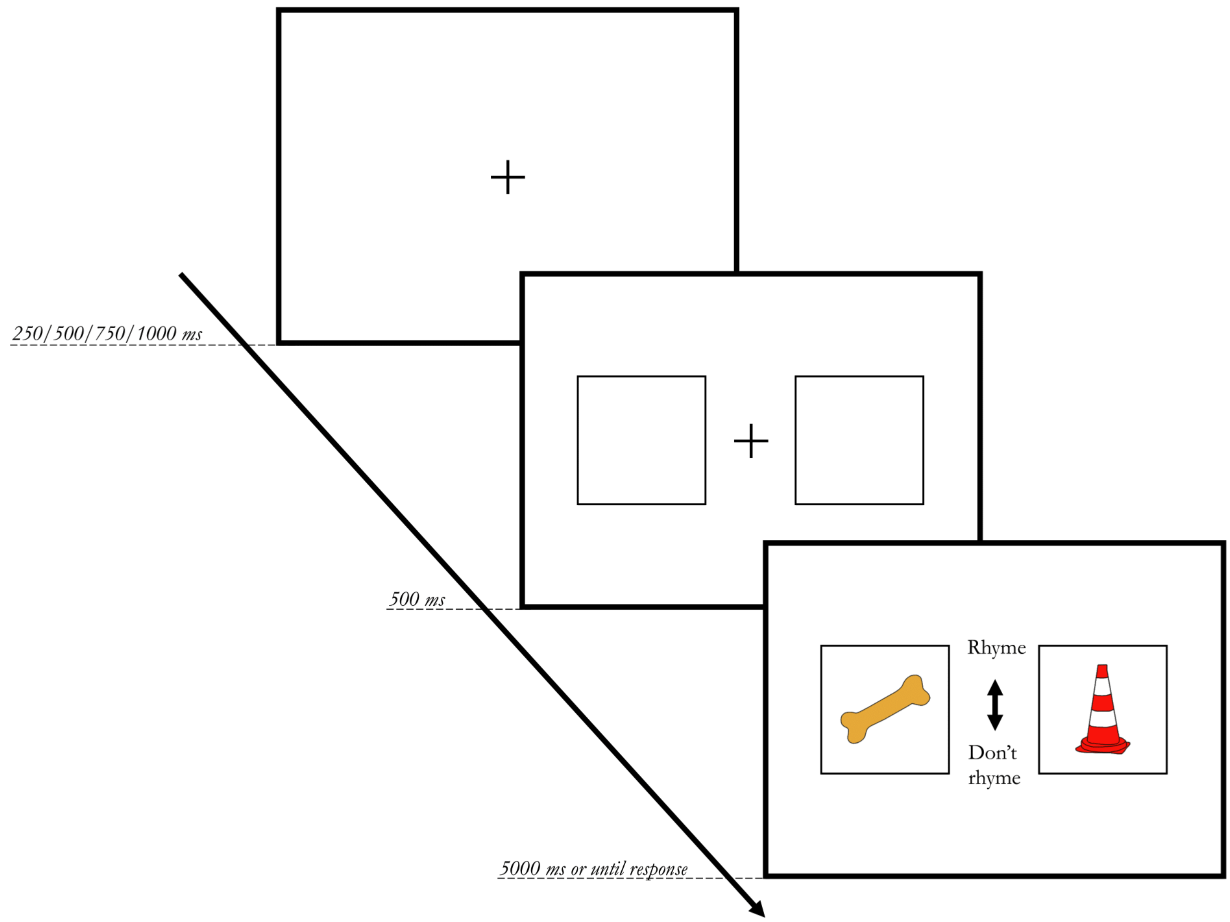

Figure 3. A sketch of a rhyme judgment trial. The stimuli here exemplify an orthographic rhyme – "bone" and "cone" – and the correct answer would therefore be "Rhyme".

2.4 Method: Task switching

2.4.1 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 4.

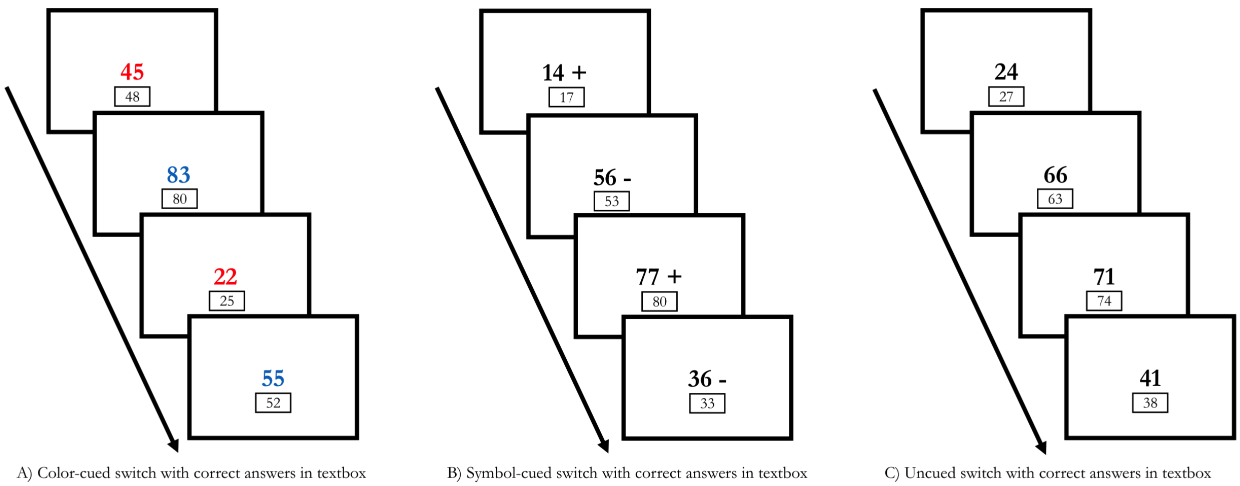

Figure 4. A sketch of the three switched conditions in the task switching experiment. Figure A shows four color-cued switch trials with correct answers, Figure B shows four symbol-cued switch trials with correct answers, and Figure C shows four un-cued switch trials with correct answers.

2.5 Method: Same/different judgments

2.5.1 Materials. This experiment used three different black silhouettes of cats 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.

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 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 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").

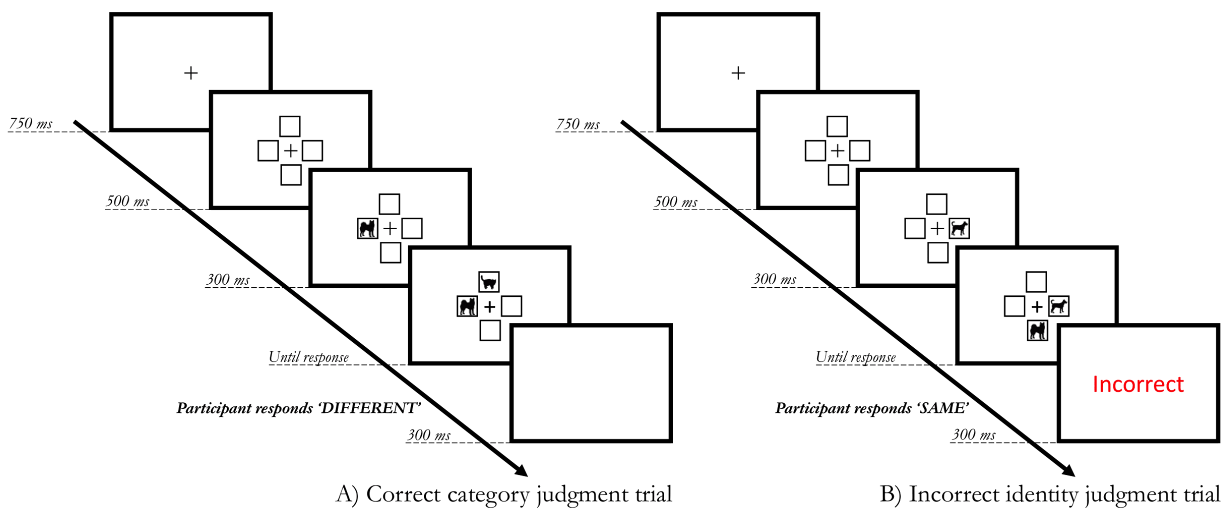

Figure 6. A sketch of the two conditions of the category judgment experiment. On Figure A, we see a correct category judgment trial where the participant responds that the cat and dog silhouettes represent different animals. On Figure B, we see an incorrect identity judgment trial where the participant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

2.6 Method: Questionnaire

After completing the four experiments, participants answered 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, Mitrenga, Wilkinson, McCarthy-Jones, & Fernyhough, 2018). See Appendix B for the full set of custom questions.

2.7 Data analysis

All analyses were conducted in R version 4.1.3 (see Appendix C for packages and citations). 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.

3 Results

3.1 Verbal working memory

3.1.1 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 7).

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

3.1.2 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, the difference was smallest for orthographically similar words (see Figure 7).

3.1.3 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 8.

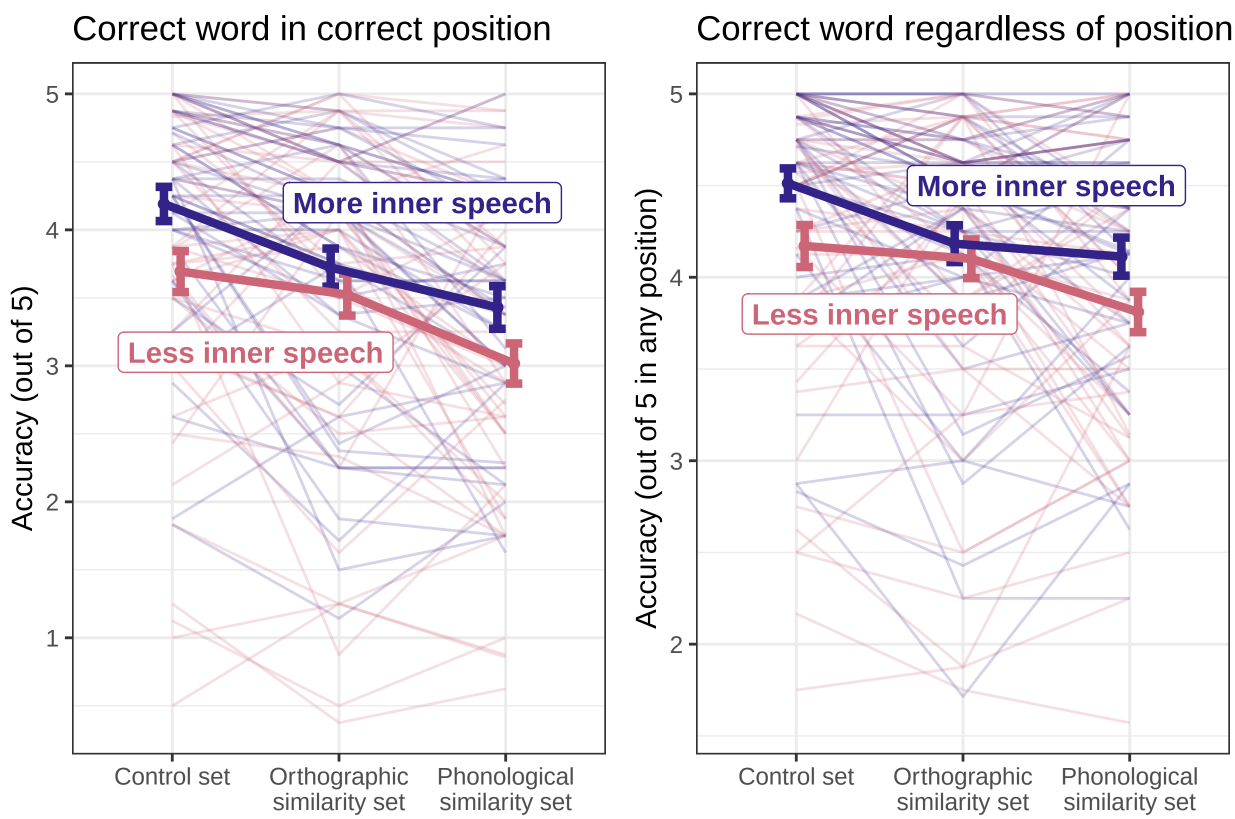

Figure 7. Score on the verbal working memory task by word set.

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).

3.2 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 = 0.05 to 1).

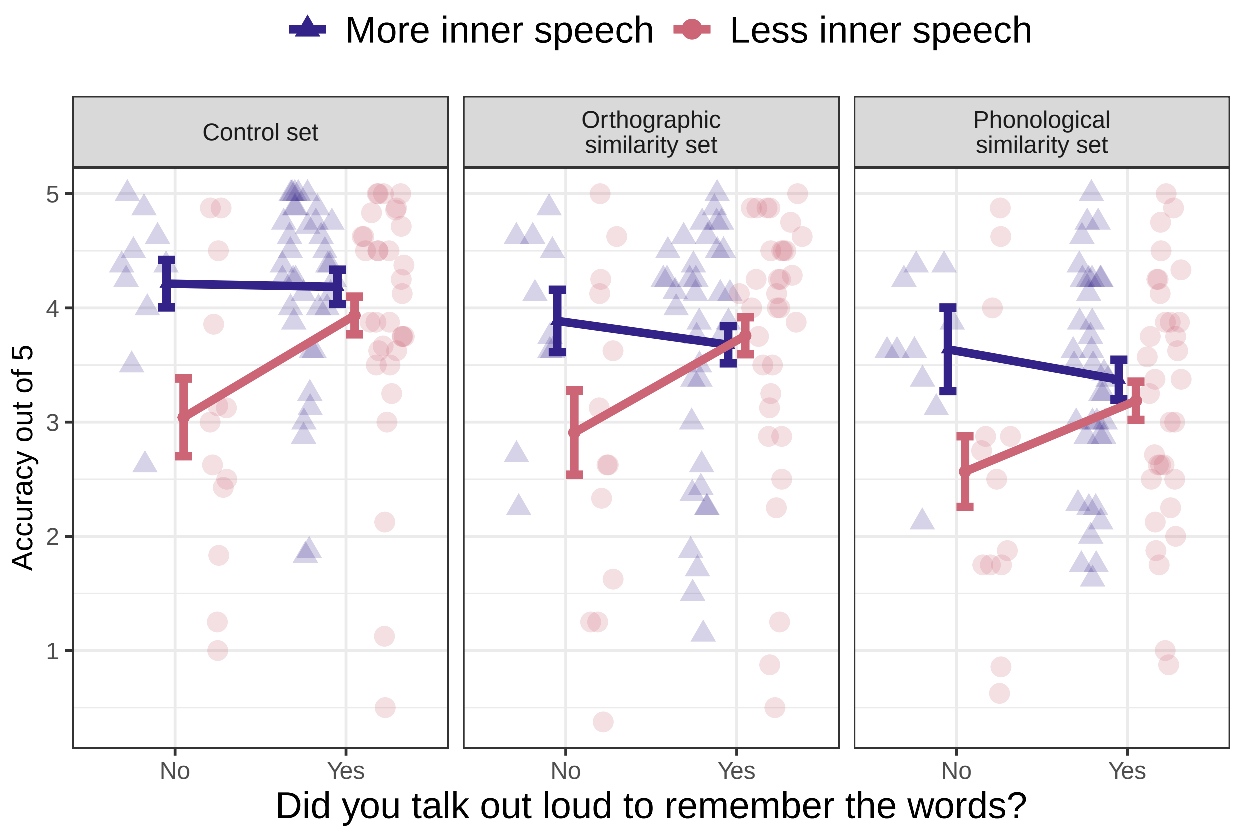

Figure 8. Verbal working memory performance by whether participants reported talking out loud to help them remember or not.

3.2.1 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 9.

3.2.2 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

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

between rhyme type and verbal score (both $p > .298$). Verbal score and name agreement also did not interact ($p > .975$).

Participants were more accurate on no-rhyme judgments (M = 95.7%) than on orthographic rhyme judgments (M = 87.5%) ($\beta = 1.30$; SE = 0.29; z = 4.49; $p < .001$) and less accurate on non-orthographic rhyme judgments (M = 79.5%) than on orthographic rhyme judgments ($\beta = -0.58$; SE = 0.26; z = -2.18; p = .029). A higher verbal score was associated with a higher likelihood of responding accurately ($\beta = 0.31$; 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).

3.2.3 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 10. Saying the words out loud diminished the

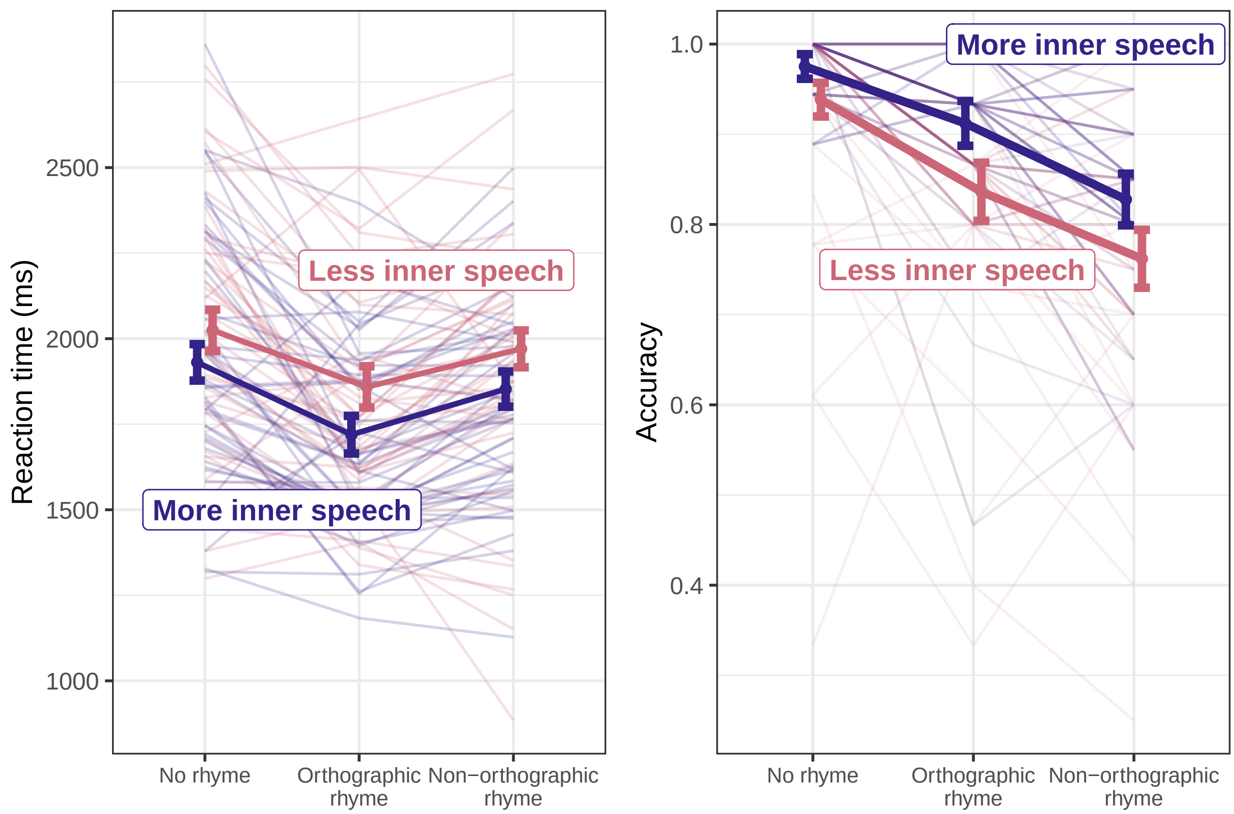

Figure 9. Reaction time and accuracy across groups by rhyme type.

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.

3.3 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.

3.3.1 Descriptive statistics: Task switching. As can be seen from Table 4 and Figure 11, accuracy was generally quite high in all conditions, and reaction times were

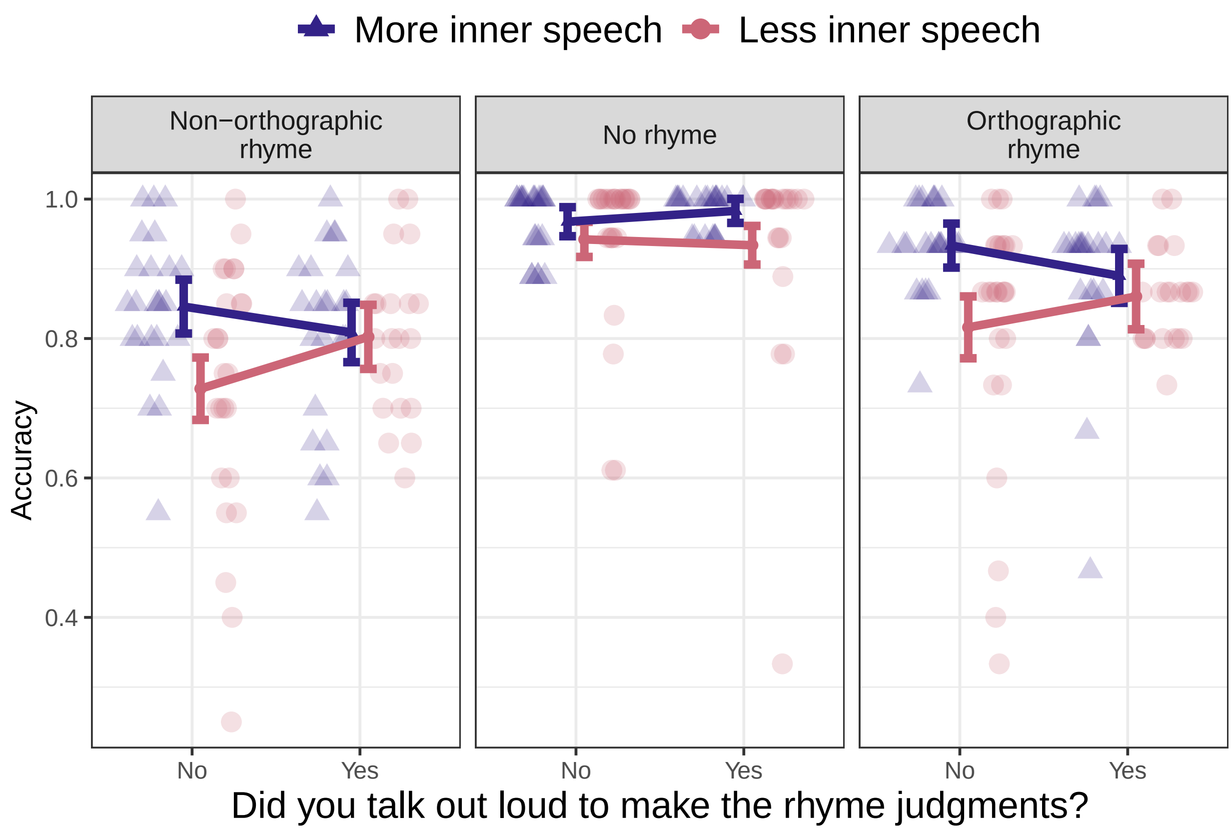

Figure 10. Reaction time and accuracy by whether participants indicated that they had talked out loud to make the rhyme judgments.

comparable across the two groups of participants.

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

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

(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) ($\beta = 0.09$; SE = 0.01; t = 8.41; $p < .001$; 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) ($\beta = 0.15$; 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 $p > .516$).

3.3.3 Strategies: Task switching. There was no significant difference between how many participants with more inner speech (20 out of 47) and how many participants

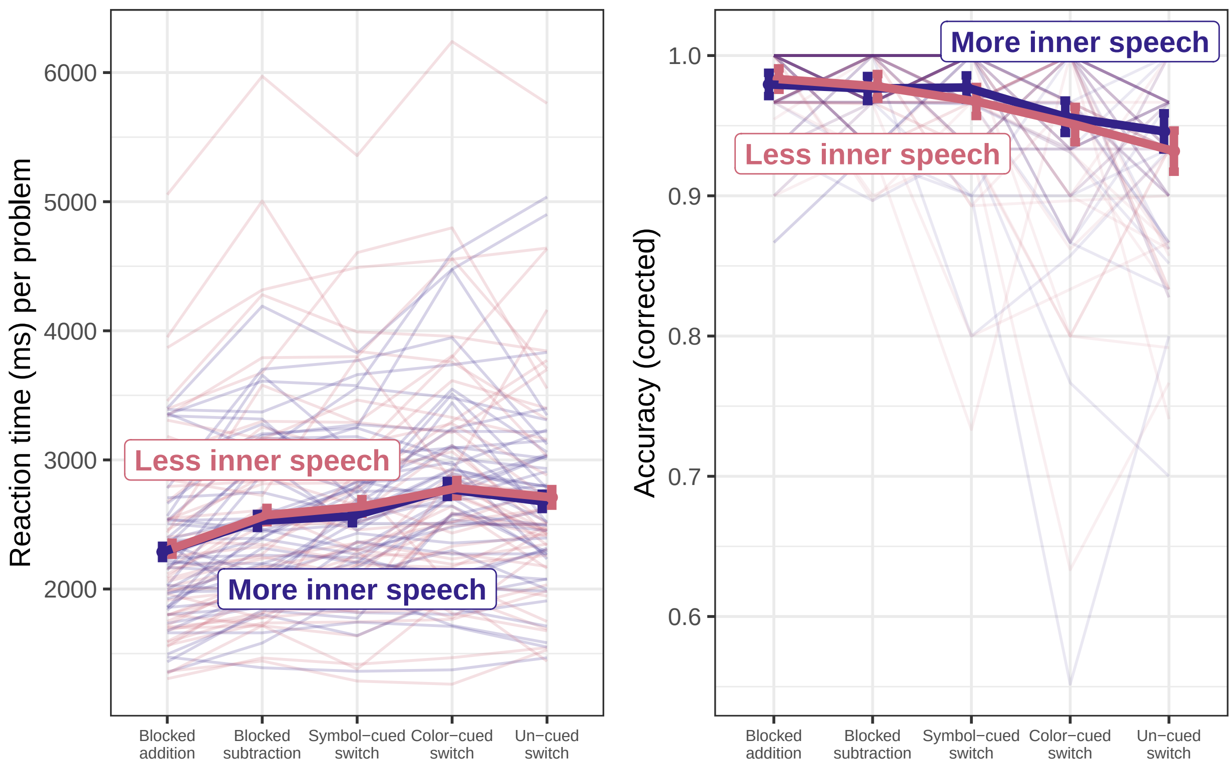

Figure 11. Reaction time and accuracy across conditions in the task switching experiment.

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 12).

3.4 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.

3.4.1 Descriptive statistics by group: Same/different judgments. See Figure 13 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

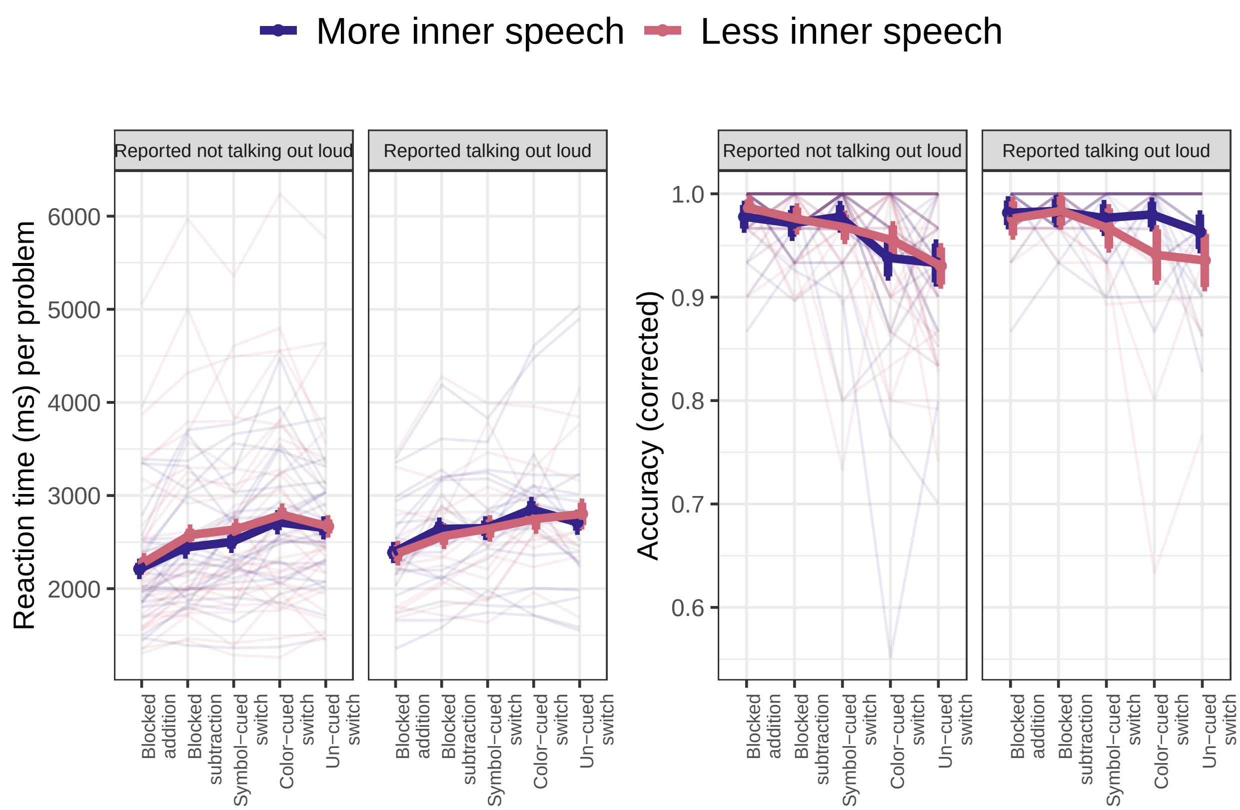

Figure 12. Reaction time (ms) and accuracy in the task switching experiment by whether participants reported talking out loud to remember the correct rule or not.

identity judgments ('are these two animals identical?').

3.4.2 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: $e^{-0.03} = 0.97$).

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 14. However, participants with more

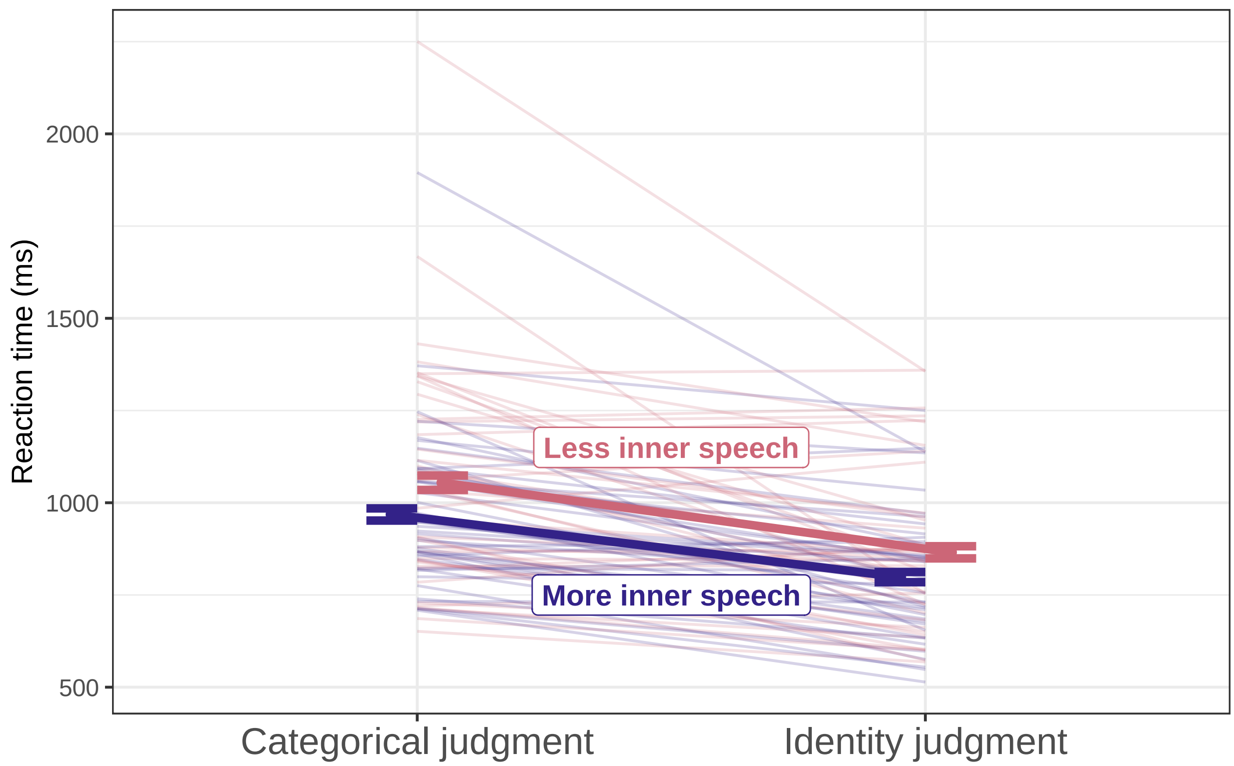

Figure 13. Reaction time in response to category or identity judgments.

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.

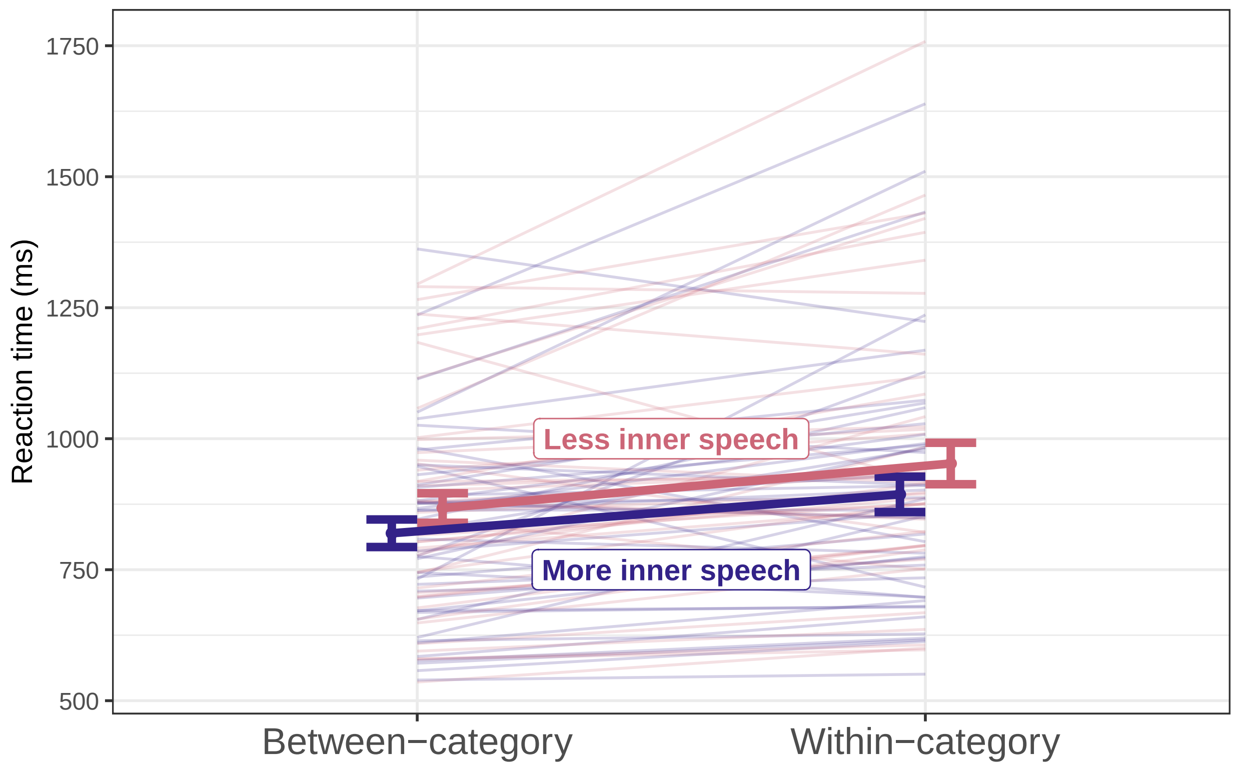

Figure 14. Reaction time on identity trials where the correct response was 'DIFFERENT' either because the two silhouettes were from different categories or different images from the same category.

3.5 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 groups responded. For reasons of space, however, we only report a few illustrative ones here (see Appendix D for plots of all the questions). The questions with the clearest differences concerned rehearsing and revising conversations where the participants with 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, our verbal representation score was mostly related to the dialogicality of inner speech (see again Figure 15) ($r(90) = .70$; $p < .001$).

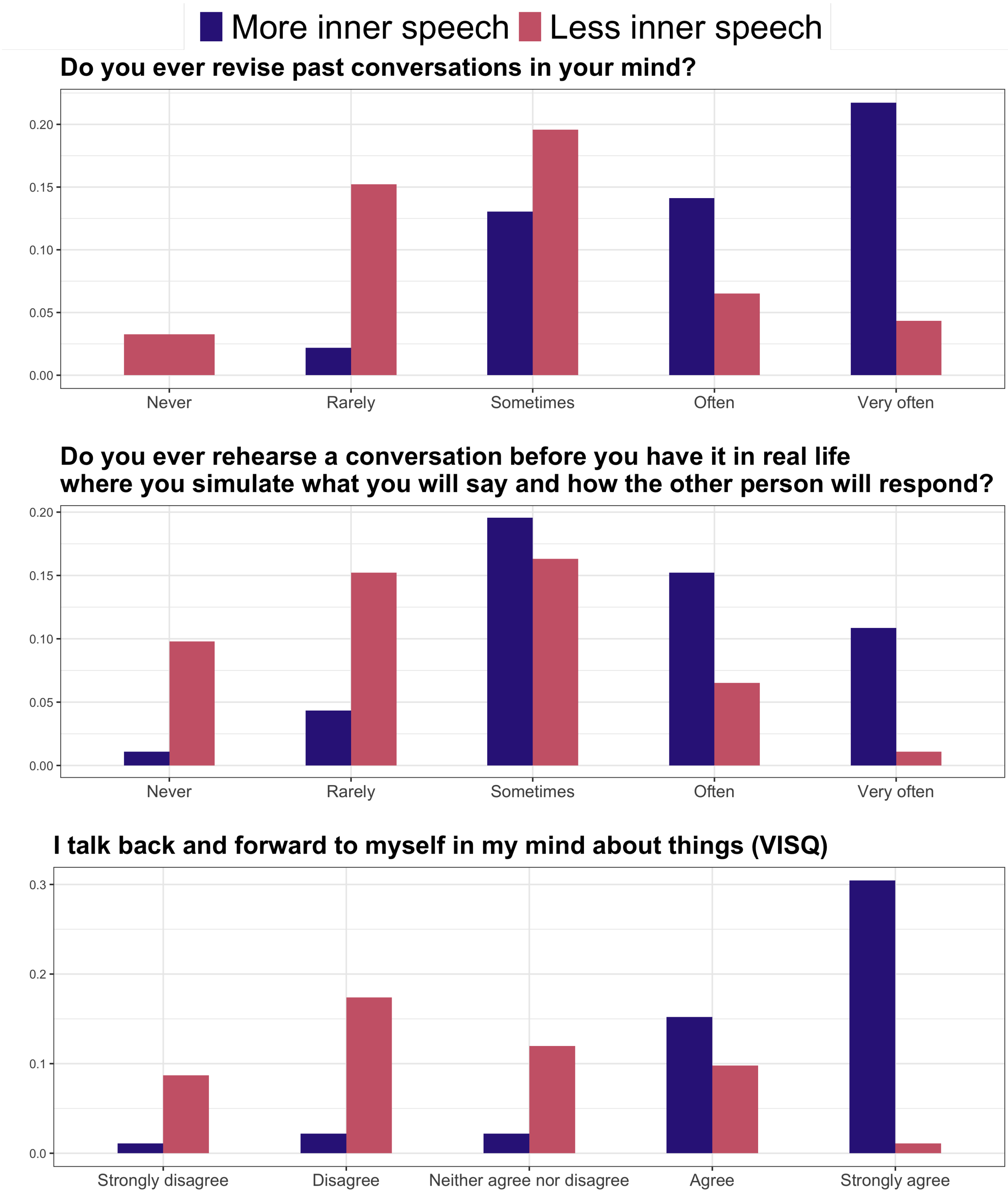

Figure 15. Grouped bar plots of proportional answers to selected custom questions concerning inner speech. Dark blue represents participants with more inner speech, and pink represents participants with less inner speech.

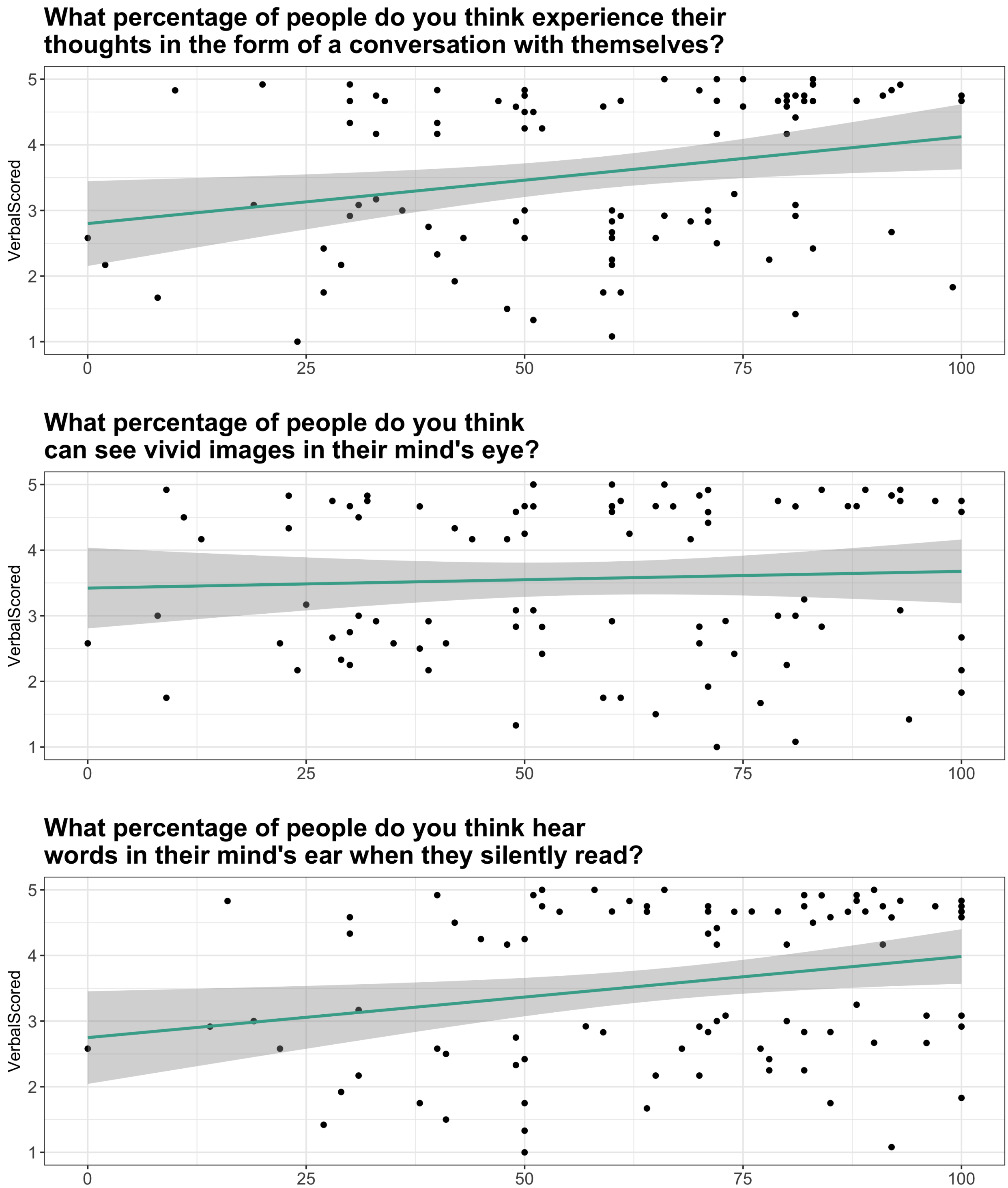

Figure 16. Scatter plots showing the correlation between verbal score on the IRQ and participants' estimates of percentages of other people with a given kind of experience.

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).

4 Discussion

Participants who report experiencing less inner speech (our sample targeted those at < 16%ile of the verbal score on the IRQ) differed in performance on several behavioral measures. First, they displayed poorer verbal working memory regardless of the material. However, contrary to our prediction, there was no indication of a weaker (or stronger) phonological similarity effect as a function of inner speech. Second, participants who report less inner speech were less accurate at judging whether the names of two images rhymed. The lack of an inner speech by nameability interaction makes it more likely that the effect stemmed from comparing phonological representations in memory rather than naming the images themselves. 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 is somewhat surprising given substantial previous research showing endogenously cued task switching being susceptible to verbal interference (Nedergaard et al., 2022). 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. In terms of our custom questionnaire, participants responded in ways consistent with their IRQ answers.

Participants with more inner speech were for example more likely to rehearse past and future conversations and to estimate that others experience an inner voice when they read and that other people experience their thoughts in the form of a conversation with themselves.

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, the term *endophasia* already exists as a term for inner speech (Bergounioux, 2001; Loevenbruck et al., 2018). The term also avoids subsuming a lack of inner speech under "aphantasia" (Monzel, Mitchell, Macpherson, Pearson, & Zeman, 2022) which we would like to avoid because inner speech is both auditory and articulatory in nature (whether it is better termed "inner hearing" or "inner speaking" is also subject to debate) and because the linguistic properties of inner speech are not reducible to phonological properties (Bermúdez, 2018; Gauker, 2018; Perrone-Bertolotti et al., 2014). For these reasons, we also do not believe the previously proposed term *anauralia* is appropriate (Hinwar & Lambert, 2021).

4.1 What have we learned about people with anendophasia?

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 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. This is especially interesting as the questions that are related to the verbal factor on the Internal Representations Questionnaire (Roebuck &

Lupyan, 2020) and which we used for participant selection are about the format of spontaneous thought (e.g., 'I think about problems in my mind in the form of a conversation with myself' and 'If I am walking somewhere by myself, I often have a silent conversation with myself'). There is some evidence that spontaneously occurring inner speech and experiment-elicited inner speech are not necessarily comparable and have different neural substrates (Hurlburt, Alderson-Day, Kühn, & Fernyhough, 2016). This makes it remarkable that our participants' reports of spontaneous inner speech seem related to their ability to use internal verbalization and verbal working memory. It is also interesting that performance was in many cases related to verbal score as a continuous factor which indicates that anendophasia is not an all-or-nothing phenomenon, much like aphantasia does not appear to be (Dance, Ipser, & Simner, 2022).

We did find evidence that using other strategies than internal verbalization could reduce 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 could be the reason why there was reduced difference in performance between the two groups for this word set. Similarly, the finding that the two groups did not differ in either reaction time or accuracy on the task switching experiment could suggest that while the inner voice can be used as a behavioral self-cue, other and equally effective strategies may be available. As mentioned in the Introduction, different strategies resulting in similar behavioural outcomes have also been found studies of people with aphantasia

(Keogh et al., 2021).

4.2 Relations to visual imagery, condensed inner speech, and unsymbolized thought

Regarding the parallels with aphantasia, it is important to note that the analogy can only take us so far. Given the findings from the present study, it seems unlikely that people with anendophasia are completely unable to verbalize internally like some people with aphantasia are completely unable to visualize. Our participants with anendophasia were not totally unable to make covert rhyme judgments, for example, as even the participants who reported not naming the images out loud performed above chance. They just found the task more difficult than the participants with more inner speech did. Instead of a total inability to verbalize internally, 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). Given that individuals with anendophasia had issues specifically with tasks that required storage and comparisons of phonological representations, it could be the case that they experience a kind of inner speech without "hearing" the speech sounds or "feeling" the articulation. Indeed, some individuals from the online communities reported that they do experience words but not the sounds of words when they think. Of the participants in the present study, one 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'. The informal reports of individuals with anendophasia thus parallel findings from Descriptive Experience Sampling of both "wordless" inner speech and unsymbolized thinking, akin to "thinking in ideas". These kinds of inner speech can potentially be usefully conceptualised as different levels of condensation of inner speech (Vicente & Martinez-Manrique, 2016). It still remains an open question whether individuals with anendophasia experience highly condensed inner speech with attenuated imagery or purely unsymbolized thought. This question could potentially be addressed through studies investigating whether or how much the categories of natural language influence individuals with anendophasia. Such study designs could for example be inspired by color categorization studies (Gilbert, Regier, Kay, & Ivry, 2006, 2008; Winawer et al., 2007).

4.3 Limitations of the present study

One limitation of our work is its reliance on wholly subjective questions for measuring inner speech. Considering that our focus is on the behavioral correlates of 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 are often wrong when they report their experience (Hurlburt & Schwitzgebel, 2011), especially if such reports take place retrospectively and require interpretations (Berger, Dennehy, Bargh, & Morsella, 2016; Ericsson & Simon, 1980; Nisbett & Wilson, 1977). It would therefore be helpful to supplement subjective assessments with physiological measures of the sort becoming possible for differences in visual imagery like investigating priming with binocular rivalry (Keogh & Pearson, 2018) or effects of visual imagery on pupil dilation (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). However, future studies could include separate measures of conscientiousness (e.g., using the Big Five Inventory, John, Donahue, & Kentle, 1991) and general intelligence, insofar as such exists (e.g., using Raven's progressive matrices, Raven, 2000).

4.4 Future directions

Just as in aphantasia, it could be the case that individual differences in inner speech remain largely undiscovered because people use alternative but equally efficient strategies for solving problems (see e.g., Keogh et al., 2021). We see some indications in our present study as well with the different effects of using a talk-out-loud strategy for the two groups. Such strategy differences should be explored in future studies, ideally through experiments where different strategies would show different behavioral profiles.

If it is correct that what people with anendophasia experience is highly condensed inner speech rather than no inner speech at all, this would also lead to predictions about the functions of their inner speech. For example, they should be less likely to use it in contexts where the specific words and sounds are important, such as as a mnemonic aid (e.g., for rehearsing a shopping list) or for simulating conversations. Indeed, the most striking difference between the two groups in the questionnaire was that participants with more inner speech spent more time rehearsing past and future conversations which makes us wonder what kind of consequences this might have. Would we expect people with more inner speech to be somehow "better" at conversations? Or maybe worse because they over-rehearse? It seems that inner speech is linked to social interactions so, in future studies, we would like to assess social cognitive abilities in populations with and without habitual inner speech. This could for example be with an adult version of the Faux Pas test (e.g., Baron-Cohen, ORiordan, Stone, Jones, & Plaisted, 1999; Thiébaut et al., 2015) where participants judge situations with social mishaps (did a faux pas occur and, if so, why was it a faux pas). This would indicate whether differences in inner speech use are related to social abilities not strictly reliant on communication.

5 Conclusion

Not everyone experiences inner speech. We proposed a name for a lack of the experience of inner speech: anendophasia. Participants with anendophasia were worse at making rhyme judgments in response to images and remembering a list of words. However, they did not differ from the control group in either task switching performance or visual discrimination judgments. They reported less auditory imagery generally (e.g., had songs stuck in their heads less often) and otherwise responded to our custom questionnaire in ways consistent with less propensity to engage in habitual inner speech. 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 due to people with anendophasia using alternative strategies. It is an open question whether anendophasia is actually a lack of inner speech or simply a lack of the experience of inner speech because of weak or absent articulatory-auditory imagery.

6 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.

Aust, F., & Barth, M. (2022). *papaja: Prepare reproducible APA journal articles with R Markdown*. Retrieved from https://github.com/crsh/papaja

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.

Baron-Cohen, S., ORiordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). *Journal of Autism and Developmental Disorders*, *29*(5), 407–418. https://doi.org/10.1023/a:1023035012436

Barth, M. (2022). *tinylabels: Lightweight variable labels*. Retrieved from https://cran.r-project.org/package=tinylabels

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., & Maechler, M. (2021). *Matrix: Sparse and dense matrix classes and methods*. Retrieved from https://CRAN.R-project.org/package=Matrix

Berger, C. C., Dennehy, T. C., Bargh, J. A., & Morsella, E. (2016). Nisbett and wilson (1977) revisited: The little that we can know and can tell. *Social Cognition*, *34*(3), 167–195.

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.

Bermúdez, J. L. (2018). Inner Speech, Determinacy, and Thinking Consciously about Thoughts. In *Inner Speech: New Voices* (Vol. 1, pp. 199–220). Oxford University Press. https://doi.org/10.1093/oso/9780198796640.003.0008

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.

Dahl, D. B., Scott, D., Roosen, C., Magnusson, A., & Swinton, J. (2019). *Xtable: Export tables to LaTeX or HTML*. Retrieved from https://CRAN.R-project.org/package=xtable

Dance, C., Ipser, A., & Simner, J. (2022). The prevalence of aphantasia (imagery weakness) in the general population. *Consciousness and Cognition*, *97*, 103243.

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.

Dowle, M., & Srinivasan, A. (2021). *Data.table: Extension of 'data.frame'*. Retrieved from https://CRAN.R-project.org/package=data.table

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.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215.

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

Fernyhough, C. (2004). Alien voices and inner dialogue: Towards a developmental account of auditory verbal hallucinations. *New Ideas in Psychology*, *22*(1), 49–68. https://doi.org/10.1016/j.newideapsych.2004.09.001

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.

Gauker, C. (2018). Inner Speech as the Internalization of Outer Speech. In *Inner Speech: New Voices* (Vol. 1, pp. 53–77). Oxford University Press. https://doi.org/10.1093/oso/9780198796640.003.0003

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.

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

Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2008). Support for lateralization of the whorf effect beyond the realm of color discrimination. *Brain and Language*, *105*(2), 91–98.

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In S. Monsell & J. Driver (Eds.), *Attention & Performance XVIII: Control of Cognitive Processes*. Cambridge, MA: MIT Press.

Grandchamp, R., Rapin, L., Perrone-Bertolotti, M., Pichat, C., Haldin, C., Cousin, E., ... Lœvenbruck, H. (2019). The ConDialInt Model: Condensation, Dialogality, and Intentionality Dimensions of Inner Speech Within a Hierarchical Predictive Control Framework. *Frontiers in Psychology*, *10*, 2019. https://doi.org/10.3389/fpsyg.2019.02019

Harrell Jr, F. E., Charles Dupont, with contributions from, & others., many. (2021). *Hmisc: Harrell miscellaneous*. Retrieved from https://CRAN.R-project.org/package=Hmisc

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

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from https://CRAN.R-project.org/package=purrr

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., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the ecological validity of thinking on demand: Neural correlates of elicited vs. Spontaneously occurring inner speech. *PLoS One*, *11*(2), e0147932.

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.

Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. *Cortex*, *105*, 61–73.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). Big five inventory. *Journal of Personality and Social Psychology*.

Kassambara, A. (2020). *Ggpubr: 'ggplot2' based publication ready plots*. Retrieved from https://CRAN.R-project.org/package=ggpubr

Kassambara, A. (2021). *Rstatix: Pipe-friendly framework for basic statistical tests*. Retrieved from https://CRAN.R-project.org/package=rstatix

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.

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex*, *105*, 53–60.

Keogh, R., Wicken, M., & Pearson, J. (2021). Visual working memory in aphantasia: Retained accuracy and capacity with a different strategy. *Cortex*, *143*, 237–253.

Kothe, E., Callegher, C. Z., Gambarota, F., Linkersdörfer, J., & Ling, M. (2021). *Trackdown: Collaborative writing and editing of r markdown (or sweave) documents in google drive*. https://doi.org/10.5281/zenodo.5167320

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

@KylePlantEmoji. (2020). Fun fact: Some people have an internal narrative and some don't. Twitter. Retrieved from https://twitter.com/KylePlantEmoji/status/1221713792913965061?s=20&t=JMap2sXmuh-XA7h5CDxZ5A

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. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, *3*, 54.

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

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.

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames*. Retrieved from https://CRAN.R-project.org/package=tibble

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

Nash, J. C. (2014). On best practice optimization methods in R. *Journal of Statistical Software*, *60*(2), 1–14. https://doi.org/10.18637/jss.v060.i02

Nash, J. C., & Varadhan, R. (2011). Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software*, *43*(9), 1–14. https://doi.org/10.18637/jss.v043.i09

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.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231.

Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: The flexible abstractness of inner speech. *Memory & Cognition*, *38*(8), 1147–1160.

Pedersen, T. L. (2021). *Ggforce: Accelerating 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=ggforce

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/

Raven, J. (2000). The raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1–48.

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.

Sarkar, D. (2008). *Lattice: Multivariate data visualization with r*. New York: Springer. Retrieved from http://lmdvr.r-forge.r-project.org

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

Terry M. Therneau, & Patricia M. Grambsch. (2000). *Modeling survival data: Extending the Cox model*. New York: Springer.

Thiébaut, F. I., White, S. J., Walsh, A., Klargaard, S. K., Wu, H.-C., Rees, G., & Burgess, P. W. (2015). Does faux pas detection in adult autism reflect differences in social cognition or decision-making abilities? *Journal of Autism and Developmental Disorders*, *46*(1), 103–112. https://doi.org/10.1007/s10803-015-2551-1

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

Vygotsky, L. S. (1962). *Thought and language*. Cambridge: Massachusetts Institute of Technology. New York; London: John Wiley.

Wei, T., & Simko, V. (2021). *R package 'corrplot': Visualization of a correlation matrix*. Retrieved from https://github.com/taiyun/corrplot

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations*. Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors)*. Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H. (2021b). *Tidyr: Tidy messy data.* Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation.* Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., Henry, L., Pedersen, T. L., Luciani, T. J., Decorde, M., & Lise, V. (2021). *Svglite: An 'SVG' graphics device.* Retrieved from https://CRAN.R-project.org/package=svglite

Wickham, H., Hester, J., & Bryan, J. (2021). *Readr: Read rectangular text data.* Retrieved from https://CRAN.R-project.org/package=readr

Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'.* Retrieved from https://CRAN.R-project.org/package=cowplot

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.

Xie, Y., & Allaire, J. (2022). *Tufte: Tufte's styles for r markdown documents.* Retrieved from https://CRAN.R-project.org/package=tufte

Zeileis, A., & Croissant, Y. (2010). Extended model formulas in R: Multiple parts and multiple responses. *Journal of Statistical Software*, *34*(1), 1–13. https://doi.org/10.18637/jss.v034.i01

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.

Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and pipe syntax.* Retrieved from https://CRAN.R-project.org/package=kableExtra

7 Supplemental materials

7.1 Appendix A: Materials for the rhyme judgment experiment

See ‘rhyming_images’ folder for the images used. See Table 5 for all image files used along with their name agreement scores from a separate experiment.

Table 5

Image files for the rhyme judgment experiment including name agreement (0 to 1) from a separate validation experiment.

File name	Name agreement
bag.png	0.95
bear.png	0.95
bed.png	0.95
beer.png	0.75
bell.png	0.90
bin.png	0.10
bone.png	0.95
boot.png	0.95
box.png	0.95
brain.png	0.95
bread.png	0.90
cab.png	0.05
cake.png	0.05
cat.png	0.90
cave.png	0.55
chain.png	0.95
chair.png	0.95
chess.png	0.25
chin.png	0.80
claw.png	0.55
clock.png	0.95
cone.png	0.70
crab.png	0.95
crane.png	0.70
dart.png	0.90
deer.png	0.65
dog.png	0.95

Table 5

Image files for the rhyme judgment experiment including name agreement (0 to 1) from a separate validation experiment. (continued)

File name	Name agreement
door.png	0.95
drawer.png	0.80
dress.png	0.85
drum.png	1.00
egg.png	0.95
eye.png	0.95
fan.png	0.75
flag.png	0.90
fly.png	0.95
fox.png	0.85
hair.png	0.90
hat.png	1.00
heart.png	0.90
house.png	0.95
jar.png	0.95
key.png	0.95
king.png	0.80
lab.png	0.70
leg.png	0.85
man.png	0.80
moon.png	0.95
mouse.png	0.85
nail.png	0.95
nose.png	1.00
park.png	0.70
pear.png	0.95
plane.png	0.35
pope.png	0.45
rake.png	1.00
ring.png	0.90
rope.png	1.00
rose.png	0.95
saw.png	1.00

Table 5

Image files for the rhyme judgment experiment including name agreement (0 to 1) from a separate validation experiment. (continued)

File name	Name agreement
screw.png	1.00
seal.png	0.75
shark.png	0.75
shell.png	0.15
shoe.png	0.85
snail.png	0.95
soap.png	0.80
sock.png	0.90
socks.png	0.90
spoon.png	1.00
square_rhyme.png	0.85
star.png	0.95
suit.png	0.95
thumb.png	0.90
tie.png	0.70
train.png	0.95
tree.png	0.85
triangle.png	1.00
wave.png	0.85
well.png	0.95
whale.png	0.95
wheel.png	0.75

7.2 Appendix B: Custom questionnaire items

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) I'm never in a position to ask questions in front of an audience (1)
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?	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)
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?	It's just like I'm hearing the conversation again. (4)
	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: - 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	No one (0%) to Everyone (100%)

7.3 Appendix C: R packages

R packages used: R (Version 4.1.3; R Core Team, 2022) and the R-packages *corrplot2021* (Wei & Simko, 2021), *cowplot* (Version 1.1.1; Wilke, 2020), *data.table* (Version 1.14.0; Dowle & Srinivasan, 2021), *dplyr* (Version 1.0.7; Wickham, François, Henry, & Müller, 2021), *forcats* (Version 0.5.1; Wickham, 2021a), *Formula* (Version 1.2.4; Zeileis & Croissant, 2010), *ggforce* (Version 0.3.3; Pedersen, 2021), *ggplot2* (Version 3.3.5; Wickham, 2016), *ggpubr* (Version 0.4.0; Kassambara, 2020), *Hmisc* (Version 4.5.0; Harrell Jr, Charles Dupont, & others., 2021), *kableExtra* (Version 1.3.4; Zhu, 2021), *lattice* (Version 0.20.45; Sarkar, 2008), *lme4* (Version 1.1.27.1; Bates, Mächler, Bolker, & Walker, 2015), *lmerTest* (Version 3.1.3; Kuznetsova, Brockhoff, & Christensen, 2017), *Matrix* (Version 1.4.0; Bates & Maechler, 2021), *optimx* (Nash, 2014; Version 2021.10.12; Nash & Varadhan, 2011), *papaja* (Version 0.1.1; Aust & Barth, 2022), *purrr* (Version 0.3.4; Henry & Wickham, 2020), *readr* (Version 2.1.1; Wickham, Hester, & Bryan, 2021), *rstatix* (Version 0.7.0; Kassambara, 2021), *stringr* (Version 1.4.0; Wickham, 2019), *survival* (Version 3.2.13; Terry M. Therneau & Patricia M. Grambsch, 2000), *svglite* (Version 2.0.0; Wickham, Henry, et al., 2021), *tibble* (Version 3.1.6; Müller & Wickham, 2021), *tidyr* (Version 1.1.4; Wickham, 2021b), *tidyverse* (Version 1.3.1; Wickham et al., 2019), *tinylabels* (Version 0.2.3; Barth, 2022), *trackdown* (Version 1.1.1; Kothe, Callegher, Gambarota, Linkersdörfer, & Ling, 2021), *tufte* (Version 0.12; Xie & Allaire, 2022), and *xtable* (Version 1.8.4; Dahl, Scott, Roosen, Magnusson, & Swinton, 2019).

7.4 Appendix D: Custom questionnaire results

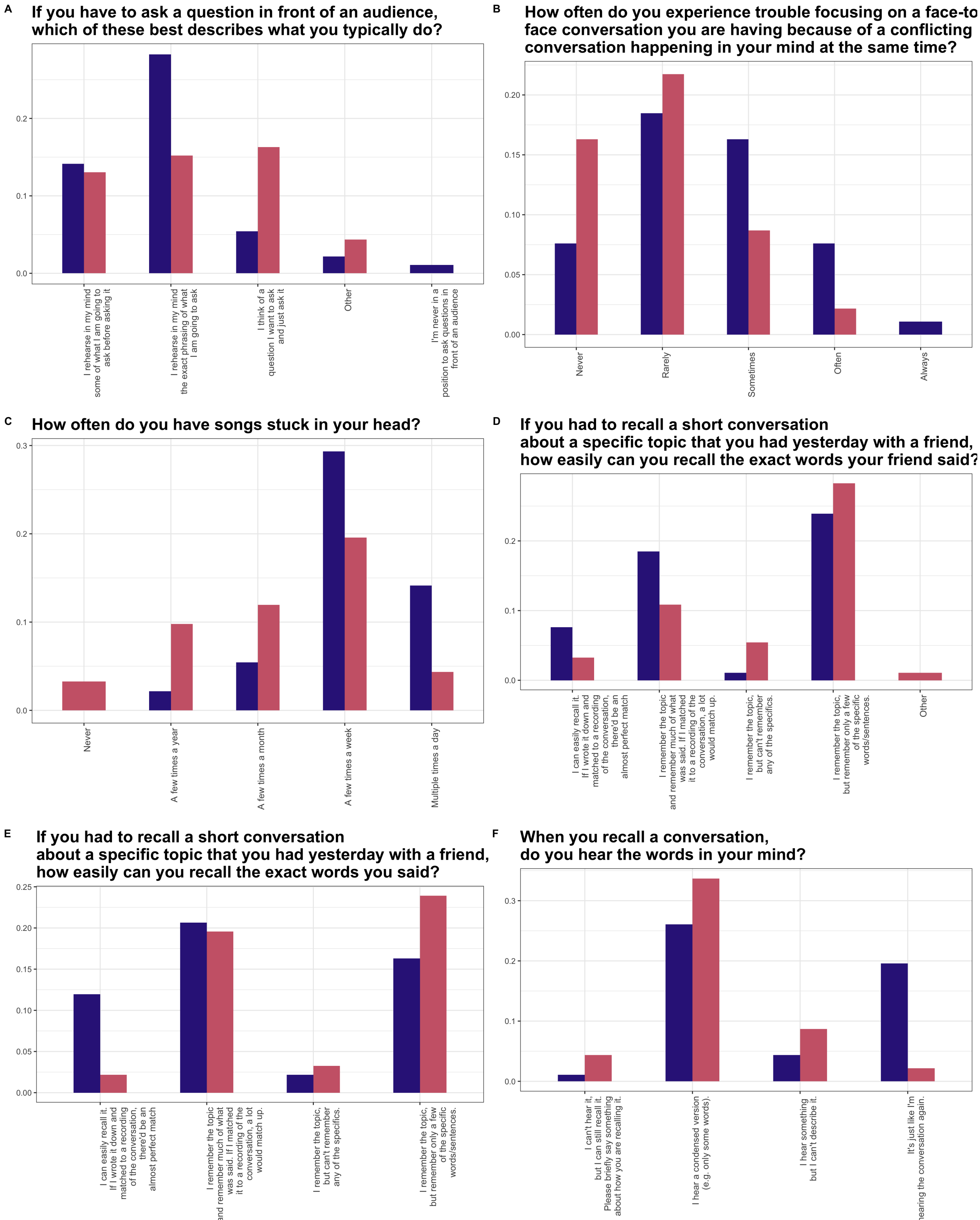

Figure 17. Grouped bar plots showing proportional answers (dark blue = more inner speech group; pink = less inner speech group).

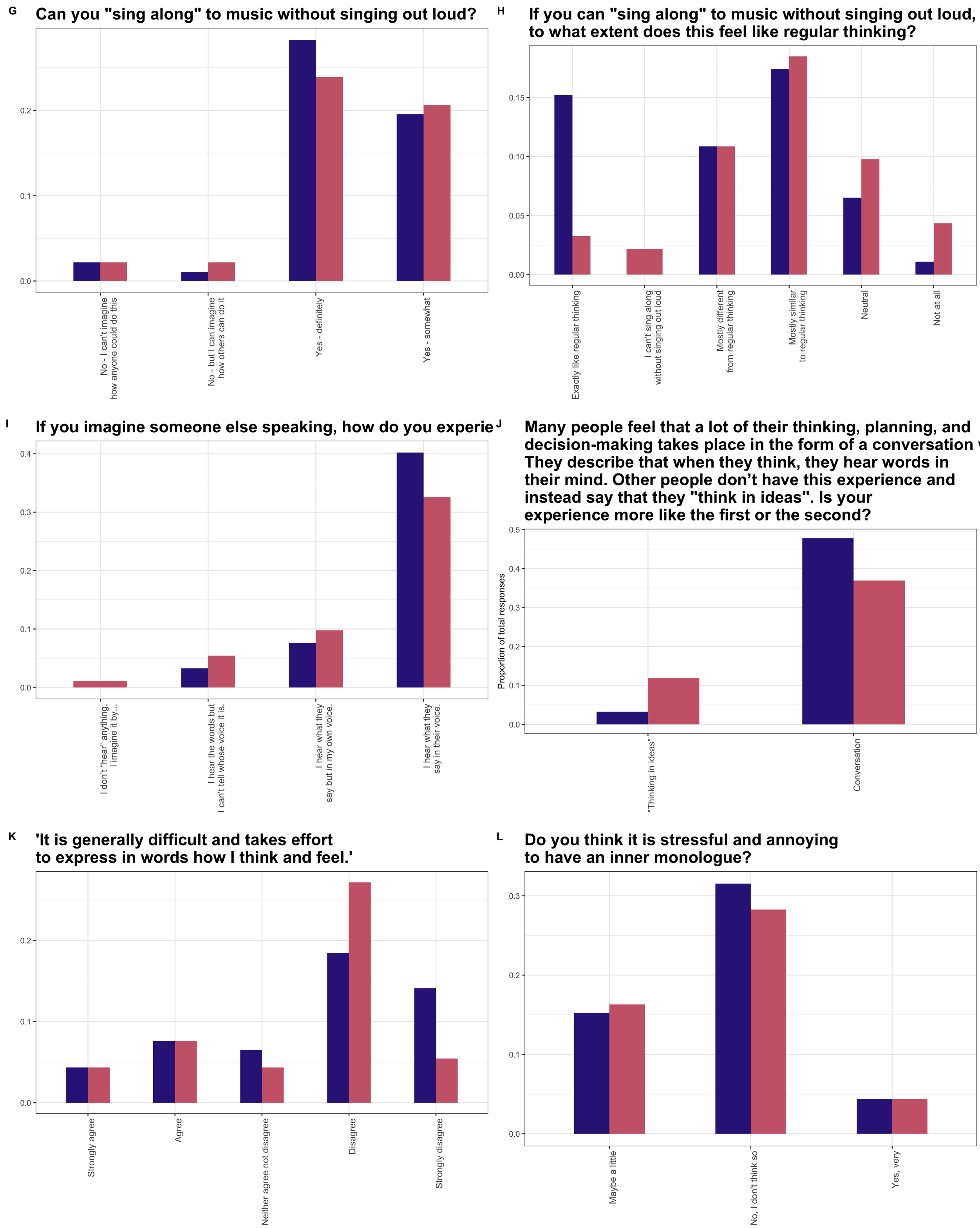

Figure 18. Grouped bar plots showing proportional answers (dark blue = more inner speech group; pink = less inner speech group).

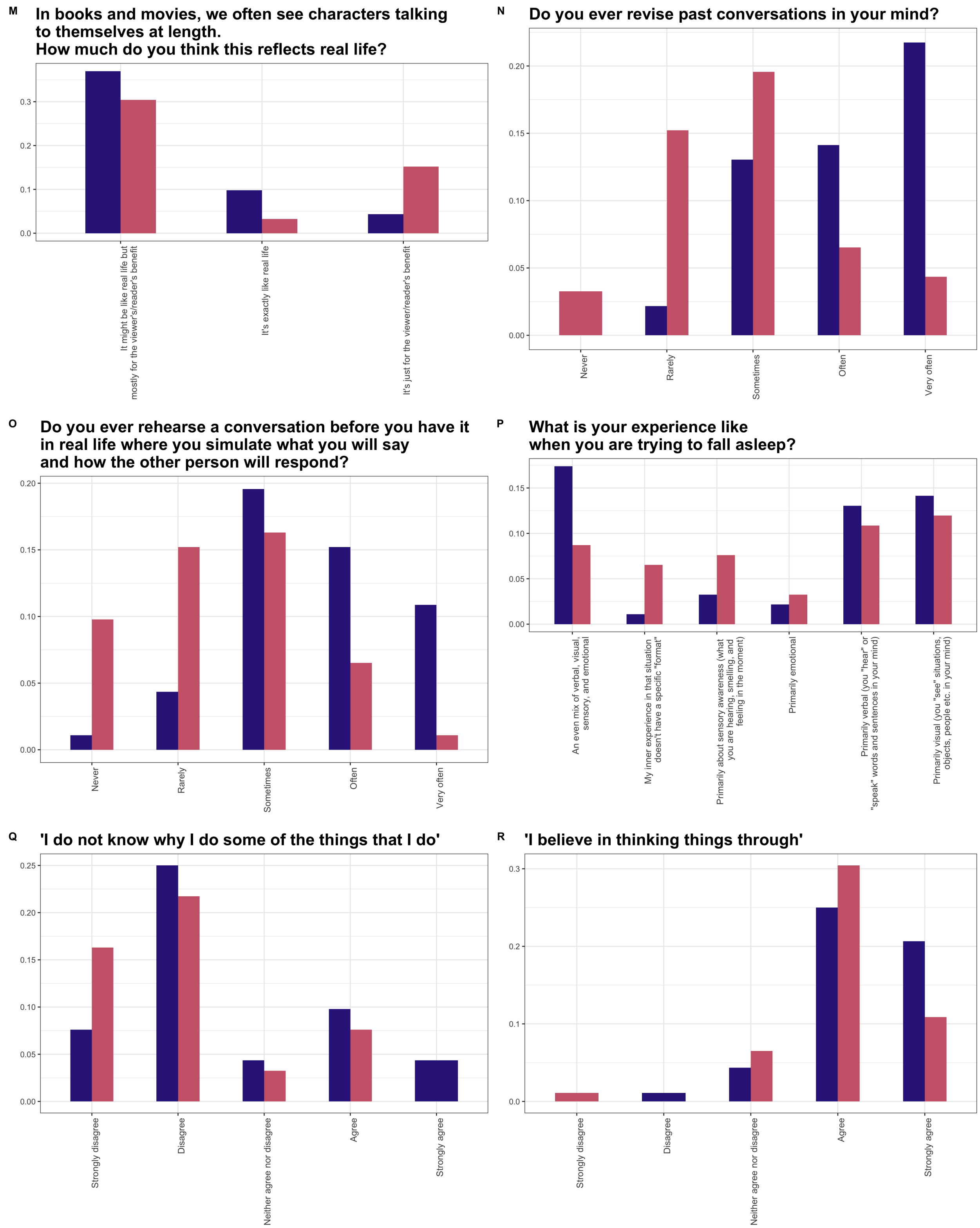

Figure 19. Grouped bar plots showing proportional answers (dark blue = more inner speech group; pink = less inner speech group).

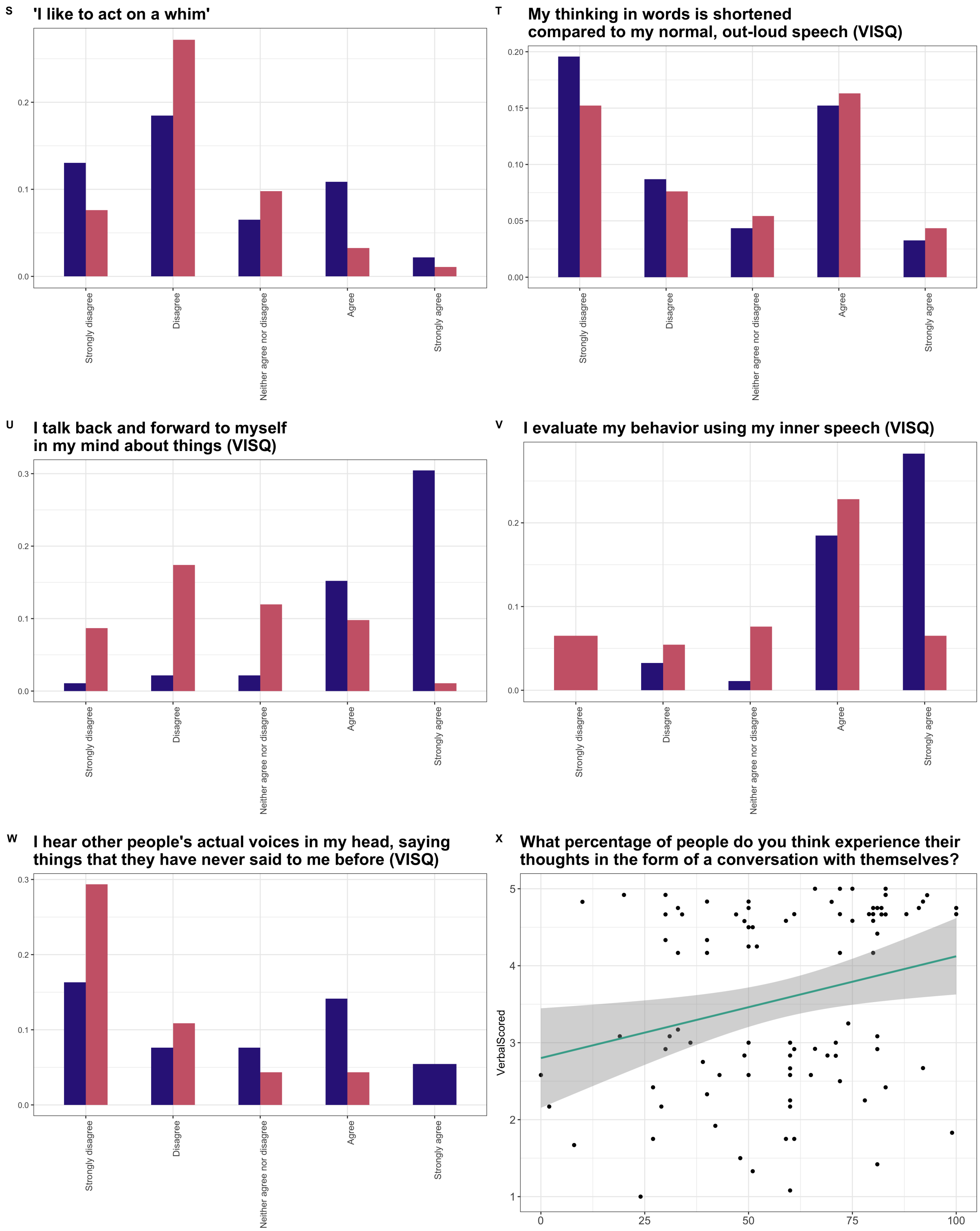

Figure 20. Grouped bar plots showing proportional answers (dark blue = more inner speech group; pink = less inner speech group).

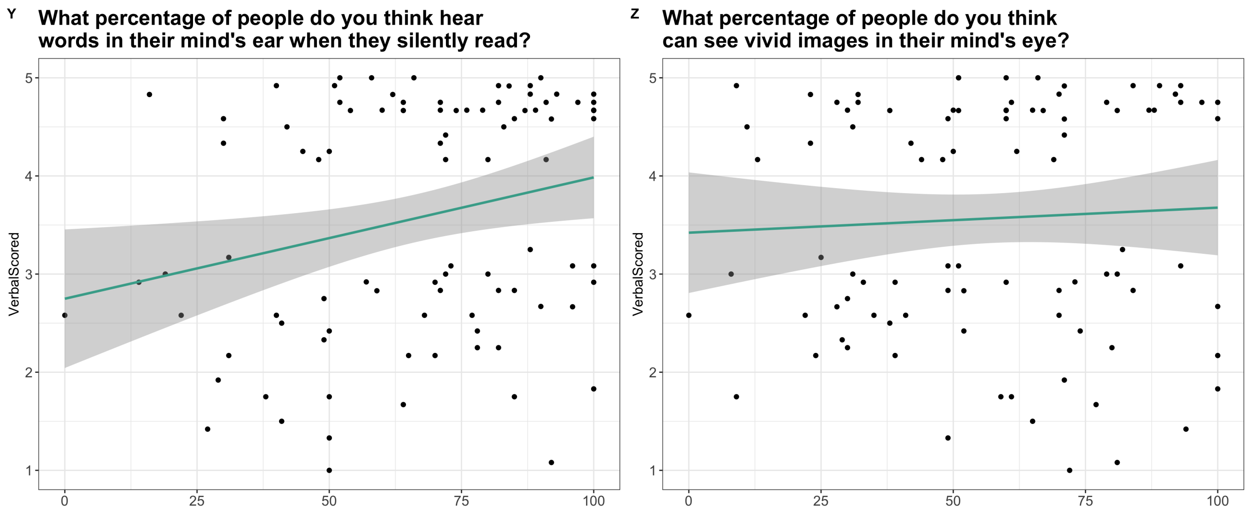

Figure 21. Figure S-W: Grouped bar plots showing proportional answers (dark blue = more inner speech group; pink = less inner speech group). Figures X-Z: Scatter plots showing correlation between verbal score on the IRQ and participants' estimates of percentages of other people with a given kind of experience.