Published in final edited form as:

Neuropsychologia. 2015 January ; 0: 121–131. doi:10.1016/j.neuropsychologia.2014.12.013.

Dopamine receptor D4 (DRD4) gene modulates the influence of informational masking on speech recognition

Zilong Xie^a, W. Todd Maddox^b, Valerie S. Knopik^{c,d}, John E. McGeary^{e,c,d}, and Bharath Chandrasekaran^a

^aDepartment of Communication Sciences & Disorders, The University of Texas at Austin, Austin TX 78712, USA

^bDepartment of Psychology, The University of Texas at Austin, Austin TX 78712, USA

^cDivision of Behavioral Genetics, Rhode Island Hospital, Providence, RI 02903, USA

^dBrown University, Providence, RI 02912, USA

^ePsychologist, Providence Veterans Affairs Medical Center, Providence, RI 02908, USA

Abstract

Listeners vary substantially in their ability to recognize speech in noisy environments. Here we examined the role of genetic variation on individual differences in speech recognition in various noise backgrounds. Background noise typically varies in the levels of energetic masking (EM) and informational masking (IM) imposed on target speech. Relative to EM, release from IM is hypothesized to place greater demand on executive function to selectively attend to target speech while ignoring competing noises. Recent evidence suggests that the long allele variant in exon III of the *DRD4* gene, primarily expressed in the prefrontal cortex, may be associated with enhanced selective attention to goal-relevant high-priority information even in the face of interference. We investigated the extent to which this polymorphism is associated with speech recognition in IM and EM conditions. In an unscreened adult sample (Experiment 1) and a larger screened replication sample (Experiment 2), we demonstrate that individuals with the *DRD4* long variant show better recognition performance in noise conditions involving significant IM, but not in EM conditions. In Experiment 2, we also obtained neuropsychological measures to assess the underlying mechanisms. Mediation analyses revealed that this *listening condition-specific* advantage was mediated by enhanced executive attention/working memory capacity in individuals with the long allele variant. These findings suggest that *DRD4* may contribute specifically to individual differences in speech recognition ability in noise conditions that place demands on executive function.

Keywords

Speech perception; Individual difference; Informational masking; Executive attention/working memory capacity; DRD4

1. Introduction

In typical social settings, speech perception often takes place in the presence of interfering background noise. Individual listeners vary substantially in their ability to perceive speech in noisy conditions (e.g., Gilbert, Tamati, & Pisoni, 2013; Song, Skoe, Banai, & Kraus, 2011; Wightman, Kistler, & O'Bryan, 2010; Wilson, McArdle, & Smith, 2007). For example, Gilbert et al (2013) showed that the overall accuracy of sentence recognition in multi-talker babble ranged from approximately 40 % to 76 % in a group of 121 young, normal-hearing adults. Previous work has examined how sensory (e.g., subcortical representation of speech sounds Chandrasekaran, Hornickel, Skoe, Nicol, & Kraus, 2009; Parbery-Clark, Strait, & Kraus, 2011; Song, et al., 2011) and cognitive factors (e.g., working memory, Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013; Koelewijn, Zekveld, Festen, Rönnerberg, & Kramer, 2012; Zekveld, Rudner, Johnsrude, & Rönnerberg, 2013) contribute to individual differences observed in speech recognition in noise tasks. A general source of individual difference is genetic variation (e.g., Bellgrove, Hawi, Kirley, Gill, & Robertson, 2005; Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Friedman et al., 2008). However, to our knowledge, no studies have examined the role of genetic factors in individual difference in speech perception in noise. To this end, the current study examined the effect of genetic variation on individual differences in executive function as it relates to speech recognition ability in challenging listening environments.

1.1 Energetic masking vs. informational masking and executive function

To recognize speech in noisy environments, one must overcome at least two types of interferences—energetic masking and informational masking (Brungart, 2001). Energetic masking (EM) occurs when noises spectro-temporally overlap with portions of target speech signals in the auditory periphery, leading to a degraded neural representation of the signals (Arbogast, Mason, & Kidd, 2002; Brungart, 2001; Freyman, Balakrishnan, & Helfer, 2004; Freyman, Helfer, McCall, & Clifton, 1999; Shinn-Cunningham, 2008). Informational masking (IM) interferes with target speech processing at more central levels of information processing. IM interference occurs even though the target signal and competing noises are relatively well represented in the auditory system (Arbogast, et al., 2002; Freyman, et al., 2004; Shinn-Cunningham, 2008). These central interferences include misattribution of components of the noise to the target (and vice versa), attentional distraction from the target, linguistic interference from the noise, and increased cognitive load (Cooke, Lecumberri, & Barker, 2008).

Previous work suggests that the mechanisms underlying EM and IM are at least partially dissociable. For example, Van Engen (2012) found that speech recognition performance in EM conditions did not predict performance in IM conditions. To cope with EM or IM, listeners are required to segregate the target source from the maskers (Shinn-Cunningham, 2008). Since the target speech and maskers are simultaneously represented in the brain (Sussman, Bregman, & Lee, 2014), to recognize target speech, listeners also need to exert top-down attention to select the target and inhibit/ignore the influences from the interfering noises (Shinn-Cunningham, 2008). As discussed before, relative to EM, IM causes more substantial central interferences, which are likely to interfere with top-down processes.

Hence, relative to EM, release from IM likely places greater demands on executive functions such as selective attention, inhibitory control, and working memory to counteract central interferences. Indeed, existing studies have shown that that working memory capacity is associated with speech recognition performance in IM conditions (Koelewijn, et al., 2012; Zekveld, et al., 2013), but not in EM conditions (Besser, Koelewijn, Zekveld, Kramer, & Festen, 2013; Koelewijn, et al., 2012; Zekveld, Rudner, Johnsrude, Heslenfeld, & Rönnerberg, 2012; Zekveld, et al., 2013). These executive processes (selective attention, inhibition, and working memory) critically depend on prefrontal cortical function (Aben, Stapert, & Blokland, 2012; Alvarez & Emory, 2006; Collette & Van der Linden, 2002; Faraco et al., 2011; Kane & Engle, 2002).

1.2 Executive function, dopamine, and dopamine D4 receptor (DRD4) gene

It is widely recognized that the neurotransmitter dopamine modulates frontostriatal circuitry critical to working memory and inhibitory control (for review, see Cools & D'Esposito, 2011; Seamans & Yang, 2004). Many studies have examined the relationship between prefrontal dopamine D1 and/or D2 receptors and prefrontal functions (e.g., Takahashi et al., 2008; Vijayraghavan, Wang, Birnbaum, Williams, & Arnsten, 2007). For example, Takahashi et al. (2008) demonstrated an inverted U-shape relation between D1 receptor expression in prefrontal cortex and executive function measured by Wisconsin Card Sorting Test. Some other studies have focused on the role of striatal dopamine in executive functions such as working memory and attention (e.g., Cools, Gibbs, Miyakawa, Jagust, & D'Esposito, 2008; Landau, Lal, O'Neil, Baker, & Jagust, 2009). For instance, Cools et al. (2008) showed that striatal dopamine synthesis capacity was positively correlated with working memory capacity as measured with listening span, with higher dopamine synthesis capacity in individuals with higher working memory capacity. Recently, there are considerable interests in understanding the role of dopamine-related genes in executive function (for review, see Barnes, Dean, Nandam, O'Connell, & Bellgrove, 2011). For example, Li et al. (2013) demonstrated that the DARPP-32 gene, which is richly expressed in the striatum, modulated auditory selective attention in situations where listeners have to focus on goal-relevant information and ignore irrelevant information.

Another well-studied dopamine gene associated with executive function is the dopamine D4 receptor (*DRD4*) gene, which is located on chromosome 11p15.5 and encodes a post-synaptic D4 dopamine receptor. Unlike DARPP-32 gene, this gene is primarily expressed in the prefrontal cortex (Oak, Oldenhof, & Van Tol, 2000). A polymorphism of *DRD4* gene lies in the 48 base pair (bp) variable number of tandem repeats (VNTR) in exon III. This polymorphism alters the sensitivity of the D4 receptor through influencing the receptor protein length in the third cytoplasmic loop (Van Tol et al., 1992). The 48-bp sequence is repeated between 2 to 11 times (Van Tol, et al., 1992). The number of repeats have been shown to associate with the potency of dopamine to inhibit cyclic adenosine monophosphate (cAMP) formation, with 7-repeat variant showing twofold reduction in the potency relative to 2- and 4- repeat (Asghari et al., 1995). Functionally, this polymorphism has been shown to associate with executive functions (e.g., Kegel & Bus, 2013), presumably via prefrontal activation (e.g., middle and inferior frontal gyrus) related to executive functions (Gilsbach et al., 2012).

In the literature, based on the repeat length, individuals have often been grouped as either “long” carriers (7 or more repeats) or “short” carriers (6 or fewer repeats). Interestingly, *DRD4* long carriers have demonstrated disrupted or enhanced executive attention (Gizer & Waldman, 2012; Kieling, Roman, Doyle, Hutz, & Rohde, 2006; Swanson et al., 2000), inhibitory control (Congdon, Lesch, & Canli, 2008; Krämer et al., 2009; Langley et al., 2004; Loo et al., 2008), and short-term memory or working memory (Altink et al., 2012; Boonstra et al., 2008; Loo, et al., 2008). To date, it remains unclear what leads to the mixed evidence regarding the role of *DRD4* in modulating executive function. A recent study suggests that *DRD4* long carriers may show enhanced selective attention to goal-relevant high-priority information even in the face of interference, but may demonstrate impaired attention to goal-irrelevant low-priority information (Gorlick et al., 2014). Of relevance to our study, this study showed that *DRD4* long carriers demonstrate superior performance on the Operation Span Task. This task measures working memory as well as domain-general executive attention (Conway et al., 2005), which requires selective attention to update and maintain high-priority items in memory while also performing a distracting secondary task. As discussed before, these executive attentional processes contribute to the release from IM. Thus, we predict that *DRD4* long carriers will demonstrate better performance in speech perception in IM conditions, but not during EM conditions.

1.3 Aims of current study

We test this hypothesis by examining the impact of the *DRD4* polymorphism on speech perception under a variety of noise conditions. In a pilot experiment (Experiment 1), with a small adult sample that was not screened for neuropsychiatric disorders, we classified participants as *DRD4* long carriers (i.e. homozygous or heterozygous for an allele of 7 or more repeats) or as *DRD4* short homozygotes (i.e. both alleles <7 repeats). We compared their sentence recognition performance in 2-talker babble (IM) and pink noise (EM) across a range of signal-to-noise ratios (SNR: -4 to 20 dB). In Experiment 2, we aimed to replicate and extend the findings from Experiment 1 with a larger independent sample that was screened for neuropsychiatric disorders. We compared sentence recognition performance in *DRD4* long and short carriers across IM and EM conditions at a fixed SNR. Importantly, we also examined the extent to which the genetic influence on speech perception was mediated via executive function by administering a battery of neuropsychological tests including measures on executive attention/working memory capacity. Consistent with a previous study demonstrating enhanced executive attentional processes in *DRD4* long carriers (Gorlick, et al., 2014), we predict that the influence of *DRD4* on speech perception in noise is routed through improved executive attention/working memory capacity.

2. Experiment 1

2.1 Material and methods

2.1.1 Participants—One hundred and thirty-one healthy adults aged 18-35 (mean $\pm$ SD: 19.05 $\pm$ 2.72; 81 female, 50 male) were recruited from the greater Austin Community. All participants completed an abbreviated version of the LEAP-Q language history questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007). All participants reported no previous history of language and hearing problems. Consistent with our previous studies

(e.g., Van Engen, Phelps, Smiljanic, & Chandrasekaran, 2014), all participants underwent a hearing screening to ensure threshold ≤ 25 dB Hearing Level (HL) at 1000 Hz, 2000 Hz and 4000 Hz for each ear. All participants provided written informed consent and received monetary compensation for their participation. All materials and procedures were approved by the Institutional Review Board at the University of Texas at Austin.

2.1.2 Genotyping—The 48-bp VNTR in the *DRD4* was assayed using previously reported methods (Hutchinson, McGeary, Smolen, Bryan, & Swift, 2002). The primer sequences used are forward, 5' AGGACCCTCATGGCCTTG-3' (fluorescently labeled), and reverse, 5'-GCGACTACGTGGTCTACTCG-3' (Lichter et al., 1993). Alleles were visualized using capillary electrophoresis. The 7 repeat allele is quite distinct from the 2–6 repeat alleles and likely originated as a rare mutational event that became more frequent as a result of positive selection (Ding et al., 2002). Participants were classified as *DRD4* long carriers (i.e. homozygous or heterozygous for an allele of 7 or more repeats) or as *DRD4* short homozygotes (i.e. both alleles <7 repeats). For quality assurance purposes in the event of ambiguity in the genotyping, the assay is run in duplicate or triplicate in order to verify the results.

2.1.3 Speech perception in noise task

Target sentences: The target stimuli consisted of 80 meaningful sentences taken from the Basic English Lexicon (Calandruccio & Smiljanic, 2012). Each sentence contained four keywords for intelligibility scoring (e.g., *The graymouseate the cheese*). One native male speaker of American English was recorded producing the full set of 80 meaningful sentences on a sound-attenuated stage at The University of Texas at Austin. The target sentences were Root mean square (RMS) amplitude equalized to 50 dB Sound Pressure Level (SPL) using Praat (Boersma & Weenink, 2010).

Maskers: Two types of maskers were used in this experiment. The first type of masker consisted of the voices of two other male talkers reciting sentences unrelated to any of the topics used in the target sentences. It was created as follows: two native male American English speakers were recorded in a sound-attenuated booth at Northwestern University as part of the Wildcat Corpus project (Van Engen et al., 2010). Each participant produced a set of 30 simple, meaningful English sentences (Bradlow & Alexander, 2007). The sentences from each talker were concatenated in random order to create 30-sentence recording without silence between sentences. These two recordings were mixed together, and then truncated to generate a 50s track of two-talker babble. The second type of masker was a 10s track of pink noise. This pink noise track was created using the Noise Generator option in Audacity (Audacity Developer Team, 2008).

Next, the two-talker babble and pink noise tracks were both equated for RMS amplitude to 54, 58, 62, 66 and 70 dB SPL using Praat (Boersma & Weenink, 2010). This created five discrete, long tracks for both masker types. Each masker track was segmented using Praat (Boersma & Weenink, 2010) to create 80 noise clips. Each noise clip was one second longer in duration than its accompanying target sentence.

Mixing targets and maskers: Each sentence was mixed with five corresponding two-talker babble clips and five corresponding pink noise clips to create stimuli of the same target sentence for each masker type with the following five signal-to-noise ratios (SNRs): -4, -8, -12, -16 and -20 dB. Each final stimulus was composed as follows: 500 ms of noise, the target and noise together, and a 500 ms noise trailer. In total, there were 400 stimuli mixed with two-talker babble (80 sentences $\times$ 5 SNRs), and 400 stimuli mixed with pink noise (80 sentences $\times$ 5 SNRs).

Testing procedure: During testing, the stimuli were bilaterally presented to participants over Sennheiser HD-280 Pro headphones. Participants were instructed that they would be listening to sentences in noise. Participants were also informed that the target sentences would always begin one-half second after the noise. In each trial, participants initiated the presentation of the stimuli by pressing a designated key on a keyboard, and were asked to type the target sentence after stimuli presentation. If they were unable to understand the entire sentence, they were asked to report any intelligible words and/or make their best guess. Participants had unlimited time to respond. There were four trials, i.e., four target sentences, in each condition and for a total of 80 trials for all 10 conditions: 2 (noise condition: two-talker babble or pink noise) $\times$ 5 (SNR: -4, -8, -12, -16, or -20 dB). The trials were presented in random order, and each one was only presented once. Responses were scored by the number of keywords correctly identified. Keywords with added or omitted morphemes were scored as incorrect.

2.1.4 Data analysis—We examine the associations of the *DRD4* VNTR polymorphism with speech perception in noise task. The data was analyzed with a linear mixed effects logistic regression where keyword identification (correct or incorrect) was the dichotomous dependent variable. Fixed effects included condition (two-talker babble or pink noise), genotype (*DRD4* long carriers or short homozygotes), SNR, and their interactions. To account for baseline differences in speech recognition performance across subjects, we included by-subject intercept as a random effect. Further, to account for the possibility that the effect of *DRD4* genotype on speech recognition may be different across subjects, we also included a random slope for each subject as a random effect (Barr, Levy, Scheepers, & Tily, 2013). Thus, the random effects were construed as this: (1 + genotype | subject). Condition and genotype were treated as categorical variables. Original SNR values (-4, -8, -12, -16, and -20) was mean-centered, and the corresponding mean-centered values were -8, -4, 0, 4 and 8. This mean-centered SNR was treated as a continuous variable. To reduce the risk of over-fitting the data, we systematically removed the insignificant fixed effects, and compared each simpler model to the more complex model using the likelihood ratio (Baayen, Davidson, & Bates, 2008). Only the results from the simplest, best-fitting model were reported in the results section. Specifically, estimates of fixed effects (i.e., β), standard errors of the corresponding estimates (i.e., *SE*), and significance of these estimates (i.e., *Z* value and *p*) were reported. The analysis was performed using the lme4 package in R (Bates, Maechler, & Bolker, 2012).

2.2 Results

2.2.1 Participants—We restricted our analysis to Caucasians whose first language was English ($N=57$, mean $\pm$ SD: 18.86 ± 2.81 , 40 female, 18 male) since previous studies have indicated poorer speech perception in noise in non-native speakers, relative to native speakers. Seven Caucasians were excluded from this sample, because of incomplete data on DNA. The final sample consisted of 50 participants. The demographics are displayed in Table 1. Results of an exact test for Hardy-Weinberg proportions using likelihood ratio (Engels, 2009) indicate that our observed genotype frequencies in the final sample significantly differ from Hardy-Weinberg equilibrium at a significant level of .05 (HWE; $p = .014$).

2.2.2 DRD4 polymorphism and Speech perception in noise—As shown in Table 2, the overall probability of correct keyword identification was higher for two-talker babble vs pink noise conditions ($p < .001$), and for the *DRD4* long carriers vs short homozygotes ($p < .001$). The effect of SNR was also significant ($p < .001$), where improving SNR increases the overall probability of correct keyword identification. Further, the results revealed a significant interaction between condition and genotype ($p = .001$). The nature of this interaction was examined by performing a second round of mixed effects logistic regressions on two-talker babble and pink noise conditions individually. The results showed that, in two-talker babble conditions, the overall probability of correct keyword identification was significantly higher for the *DRD4* long carriers than for the short homozygotes (see Fig. 1A), $\beta = 0.51$, $SE = 0.17$, $Z = 2.99$, $p = .003$. However, in pink noise conditions, the overall probability of correct keyword identification did not significantly differ between *DRD4* long carriers and short homozygotes (see Fig. 1B), $\beta = 0.06$, $SE = 0.08$, $Z = 0.65$, $p = .513$.

The results also revealed a significant interaction between condition and SNR ($p < .001$). The nature of this interaction was examined by performing a second round of mixed effects logistic regressions on two-talker babble and pink noise conditions individually. The results showed that, although lowering SNR decreases the overall probability of correct keyword identification in both two-talker babble and pink noise conditions, the intelligibility drop from SNR decrement was greater in pink noise conditions ($\beta = 0.40$, $SE = 0.01$, $Z = 32.97$, $p < .001$) than in two-talker babble conditions ($\beta = 0.24$, $SE = 0.008$, $Z = 29.23$, $p < .001$).

2.3 Discussion

Results from Experiment 1 demonstrate that the long variant of the *DRD4* gene was significantly associated with better recognition performance in noise condition that has significant informational masking (2-talker babble). We did not observe differences in recognition performance between *DRD4* long carriers and short carriers when the noise condition was primarily energetic (pink noise). These results provide preliminary evidence in support of a *listening condition-specific* advantage of the *DRD4* long allele in conditions with significant informational masking. Nonetheless, candidate gene studies have been criticized for poor replicability (Ioannidis, Ntzani, Trikalinos, & Contopoulos-Ioannidis, 2001). Hence, to increase confidence in the validity of these results, it is necessary to replicate the results in a separate study. Moreover, as the sample in Experiment 1 was not

screened for neuropsychiatric disorders, we wanted to determine if the findings still hold in a larger sample that was more systematically screened for neuropsychiatric disorders. Finally, in experiment 2 we sought to examine the mechanisms underlying the listening condition-specific advantage of *DRD4* long allele in informational masking. As proposed in the introduction section, executive attention/working memory capacity potentially underlies this listening condition-specific advantage of the *DRD4* long allele in informational masking. Hence, in experiment 2, we tested speech recognition in a screened sample across a wider range of noise conditions that can be categorized as predominantly informational (1-talker, 2-talker) or energetic (8-talker and speech-shaped noise). We also administered a battery of neuropsychological tests including measures on executive attention/working memory capacity to assess the extent to which these measures mediate the *DRD4* long allele advantage in informational masking conditions.

3. Experiment 2

3.1 Material and methods

3.1.1 Participants—This experiment is part of an ongoing, large-scale genetics and cognition project at The University of Texas at Austin. Two hundred and twenty-two healthy adults aged 18-35 (mean $\pm$ SD: 25.12 $\pm$ 4.36; 125 female, 97 male) were recruited from the greater Austin Community. All participants were screened using the Mini International Neuropsychiatric Interview (MINI) (Lecrubier et al., 1997; Sheehan et al., 1997) to ensure that they did not meet criteria for a current or past psychiatric diagnosis. Based on MINI screening results, none of the participants were taking psychoactive medication or psychotherapy at the time of the study. None of the participants reported previous history of brain trauma. All participants completed an abbreviated version of the LEAP-Q language history questionnaire (Marian, et al., 2007). All participants reported no previous history of language and hearing problems. All participants completed a battery of neuropsychological tests on various aspects about executive function (see section 4 in *Material and methods* for details). Consistent with previous studies (e.g., Van Engen, et al., 2014), all participants underwent a hearing screening to ensure threshold $\leq$ 25 dB HL at 1000 Hz, 2000 Hz and 4000 Hz for each ear. Although thresholds (and oto-acoustic emissions) may have provided a more sophisticated profile of hearing acuity, our extensive neuropsychological screening battery, cognitive tests, experiments, as well as collecting a genetic screen did not allow enough time to develop a more comprehensive hearing profile for the participants. All participants provided written informed consent and received monetary compensation for their participation. All materials and procedures were approved by the Institutional Review Board at the University of Texas at Austin.

3.1.2 Genotyping—Genotyping methods were identical to Experiment 1.

3.1.3 Speech perception in noise task

Target sentences: Sentences from the Revised Bamford-Kowal-Bench (BKB) Standard Sentence Test (Bamford & Wilson, 1979) were recorded by a female native speaker of American English in a sound-attenuated booth at Northwestern University (Van Engen,

2012). Four BKB lists were recorded, and each list contained 16 sentences and a total of 50 keywords for scoring. All sentence recordings were equalized for RMS amplitude.

Maskers: Four masker types ranging from primarily informational to primarily energetic were used in the current experiment: 1-talker babble, 2-talker babble, 8-talker babble, and speech-spectrum noise (SSN). The 1- and 2-talker babble produce primarily IM, as the confusability (and/or perceptual similarity) between these maskers and target speech is largest (Freyman, et al., 2004). SSN is the end product of the summation of an infinite number of talkers that produces mainly EM (Simpson & Cooke, 2005). The 8-talker babble produces the same amount of EM as SSN (Brungart, Chang, Simpson, & Wang, 2009), and creates almost no linguistic inference—one of the critical components of IM (Freyman, et al., 2004). These maskers were created as follows: first, eight female speakers of American English were recorded in a sound-attenuated booth at Northwestern University (Van Engen, et al., 2010). Each participant produced 30 simple, meaningful English sentences (Bradlow & Alexander, 2007). For each talker, these sentences were equalized for RMS amplitude and then concatenated to create 30-sentence strings without silence between sentences. One of these recordings was used as the 1-talker babble track. To generate 2-talker babble, the recording from a second talker was mixed with the first in Audacity (Audacity Developer Team, 2008). Six more talkers were added to create 8-talker babble. In order to generate SSN, steady-state white (i.e., flat spectrum) noise was filtered so that its spectrum matched the long-term average spectrum of the full set of 240 sentences. All masker tracks were truncated to 50 s and equated for RMS amplitude.

Mixing targets and maskers: Each target sentence was mixed with a random sample of noise such that each stimulus was composed as follows: 400ms of silence, 500 ms of noise, the target and noise together, and a 500 ms noise trailer. The signal-to-noise ratio was set to –5 dB. This ratio was chosen on the basis of a previous study (Chandrasekaran, Van Engen, Xie, Beevers, & Maddox, in press) in order to avoid floor and ceiling performances across noise conditions.

Testing procedure: The test environments, instructions, trial design and procedure, and intelligibility scoring rules were identical to Experiment 1, except that the total trials added up to 64.

3.1.4 Neuropsychological tests

Operation Span task (OSPAN): This task was used to measure working memory as well as domain-general measure of executive attention (Conway, et al., 2005). In this task, participants' primary goal was to remember a sequence of letters presented on a computer screen while performing a secondary arithmetic problem. After each sequence, participants were instructed to recall letters in the same order as they were presented and report them using the keyboard. Meanwhile, they must maintain an accuracy of 85% or above on the math problems. The task consisted of 15 recall sequences with sequence length ranging from 3 to 7 letters. Participants recalled a total of three sequences for each sequence length (3 to 7 letters) for a total of 75 letters. During scoring, correctly recalling a sequence added the length of that sequence to one's score (Unsworth & Engle, 2005). For example, correctly

recalling a sequence of six letters adds six points to one's score. Correctly recalling a sequence of seven letters adds seven more points to one's score. Meanwhile, for those sequences that participant incorrectly recalled even one letter, zero point would be added to one's score. The summed scores from all the 15 sequences served as individual's OSPAN score.

Forward Digit Span task: Forward Digit Span task was used to assess the participants' verbal short term memory (Wechsler, 1997). In this task, participants listened to a sequence of digits at a rate of 1 digit per second, and were instructed to restate the digits in the same order as they were presented. The participants listened to two sequences of each sequence length, and as long as they correctly reproduced at least one sequence of the two, they moved on to the next sequence length, for a maximum length of nine digits. The task terminated when the participants failed to reproduce two consecutive sequences at any given sequence length. Individual's Forward Digit Span score was calculated as the total number of correct sequences.

Stroop test: This task was used to measure selective attention/inhibitory control ability (Stroop, 1935). In this task, there are three conditions: word-only, color-only, and word-color. In the word-only condition, participants were asked to read a list of black-font color words (e.g., "blue") as quickly as they can. In the color-only condition, participants were asked to state the color of X's printed in different colors as quickly as they can. In the word-color condition, participants viewed a list of color words, printed in a non-matching color. They were instructed to state the color of the printed ink for the color words as quickly as they can. Each condition lasted 45 seconds, and interference scores were calculated as the difference between the word-only and the color-only conditions. The interference score for each participant was then converted to standardized Z-scores using the standard age-appropriate published norms.

3.1.5 Data analysis—First, we examined the effects of the *DRD4* polymorphism on performance in the speech perception in noise task. This analysis was identical to Experiment 1, except that we grouped the four noise conditions into two levels: primarily informational masking (1- and 2- talker babble) or primarily energetic masking (8-talker babble and SSN), and these two new levels composed of the condition variable.

Second, we assessed the relation between neuropsychological tests and speech perception in noise using Spearman correlations. The speech perception in noise performance was calculated as the percentage of correct identified keywords in the primarily informational masking (1- and 2- talker babble) or primarily energetic masking (8-talker babble and SSN) conditions separately.

Finally, we examined whether OSPAN mediates the effects of *DRD4* polymorphism on speech perception in primarily informational masking conditions following these steps (Baron & Kenny, 1986): (a) the independent variable (*DRD4* genotype) relates to the dependent variable (recognition performance in informational masking conditions); (b) the independent variable (*DRD4* genotype) relates to the mediator (OSPAN); (c) the mediator (OSPAN) relates to the dependent variable (recognition performance in informational

masking conditions); (d) when the mediator is held constant, the independent variable does not have an effect on the dependent variable (full mediation) or the relation becomes significantly smaller (partial mediation); and (e) the indirect effect of the independent variable on the dependent variable, using the Sobel test, should be significant.

3.2 Results

3.2.1 Participants—As in Experiment 1, we restricted our analysis to Caucasians whose first language was English ($N=124$, mean $\pm$ SD: 25.9 ± 4.37 , 63 female, 61 male). Nineteen Caucasians were excluded from this sample, because of incomplete data on DNA and/or OSPAN. One more participant was excluded because they did not maintain an accuracy level of at least 85% on the arithmetic problems on the OSPAN. The final sample consisted of 104 participants. The demographics are displayed in Table 3. Results of an exact test for Hardy-Weinberg proportions using likelihood ratio (Engels, 2009) indicate that our observed genotype frequencies in this screened sample differ significantly from Hardy-Weinberg equilibrium (HWE; $p < .001$).

3.2.2 DRD4 polymorphism and Speech perception in noise—As shown in Table 4, the overall probability of correct keyword identification was higher for energetic masking vs informational masking conditions ($p < .001$), and for the *DRD4* long carriers vs short homozygotes ($p < .001$). The results also revealed a significant interaction between condition and *DRD4* genotype ($p < .001$). The nature of this interaction was examined by performing a second round of mixed effects logistic regressions on informational and energetic masking conditions individually. The results showed that, in the informational masking condition, the overall probability of correct keyword identification was significantly higher for *DRD4* long carriers than for short homozygotes (see Fig. 2), $\beta = 0.57$, $SE = 0.22$, $Z = 2.59$, $p = .010$. However, in the energetic masking condition, the overall probability of correct keyword identification did not significantly differ between the *DRD4* long carriers and short homozygotes (see Fig. 2), $\beta = 0.05$, $SE = 0.11$, $Z = 0.51$, $p = .608$ ^{1,2}.

3.2.3 Relationship between neuropsychological tests and speech perception in noise—Spearman correlation analysis showed that OSPAN scores significantly correlated with speech recognition performance in informational masking conditions ($r = .274$, $p < .001$), but not in energetic masking conditions ($r = .067$, $p = .338$). Both Forward

¹We also ran the same analysis with only 1-talker babble and SSN conditions which can be considered as the logical extremes of IM and EM. The results show the same pattern. Specifically, there was a significant interaction between condition and *DRD4* genotype, $\beta = -0.93$, $SE = 0.13$, $Z = -7.45$, $p < .001$. We examined the nature of this interaction by performing a second round of mixed effects logistic regressions on 1-talker babble and SSN conditions individually. The results showed that, in 1-talker babble conditions, the overall probability of correct keyword identification was significantly higher for *DRD4* long carriers than for short homozygotes, $\beta = 1.01$, $SE = 0.41$, $Z = 2.47$, $p = .014$. However, in SSN conditions, the overall probability of correct keyword identification did not significantly differ between *DRD4* long carriers and short homozygotes, $\beta = -0.02$, $SE = 0.16$, $Z = -0.10$, $p = .920$.

²From the 84 *DRD4* short homozygotes, we selected a subgroup ($n = 20$) matched for age and sex with the *DRD4* long carriers ($n = 20$). The matched short homozygotes group was randomly selected by a research assistant who was blind to participants' performance. We ran the same analysis on speech perception in primarily informational masking (1- and 2-talker babble) or primarily energetic masking (8-talker babble and SSN). The pattern of *DRD4* polymorphism effects results still holds. Specifically, there was a significant interaction between condition and *DRD4* genotype, $\beta = -0.63$, $SE = 0.10$, $Z = -6.37$, $p < .001$. We examined the nature of this interaction by performing a second round of mixed effects logistic regressions on informational and energetic masking conditions individually. The results showed that, in informational masking conditions, the overall probability of correct keyword identification was significantly higher for *DRD4* long carriers than for short homozygotes, $\beta = 0.79$, $SE = 0.31$, $Z = 2.52$, $p = .012$. However, in energetic masking conditions, the overall probability of correct keyword identification did not significantly differ between *DRD4* long carriers and short homozygotes, $\beta = 0.16$, $SE = 0.15$, $Z = 1.06$, $p = .289$.

Digit Span and Stroop scores were not significantly associated with speech recognition performance in informational masking conditions (Forward Digit Span, $r = .087$, $p = .212$; Stroop, $r = .059$, $p = .398$), or in energetic masking conditions (Forward Digit Span, $r = -.013$, $p = .854$; Stroop, $r = .020$, $p = .778$).

3.2.4 Mediation—As results section 3 showed that only OSPAN was significantly associated with speech recognition performance in informational masking (IM) condition, we focused the mediation analysis on OSPAN. Table 5 displays the results for testing OSPAN as a mediator between *DRD4* genotype and speech recognition performance in IM conditions. The first step in the analysis showed that *DRD4* genotype was a significant predictor of recognition performance in IM conditions ($p = .011$), which explained variance of 2.6%. The second step in the analysis demonstrated that *DRD4* genotype was a significant predictor of OSPAN ($p = .017$), which explained variance of 5%. The third step in the analysis showed that OSPAN was a significant predictor of performance in IM conditions ($p < .001$), which explained variance of 6.5%. The fourth step in the analysis aimed at testing the model in which *DRD4* genotype and OSPAN were entered simultaneously. *DRD4* genotype was no longer a significant predictor of performance in IM conditions ($p = .404$), but OSPAN remained a significant predictor of performance in IM conditions ($p = .001$). And the interaction between *DRD4* genotype and OSPAN were not significant predictors of performance in IM conditions ($p = .714$). This model had an explained variance of 7%.

In the final step, the Sobel test of the indirect relation between *DRD4* genotype and speech recognition performance in IM conditions was significant ($p = .039$). In sum, the model provides evidence that OSPAN is an almost complete mediator of the relation between *DRD4* and speech recognition performance in IM conditions.

4. General discussion

4.1 Summary of findings

The goal of the current study was to examine the effect of genetic variation on individual differences in executive function as it relates to speech recognition ability under various challenging listening environments. Specifically, we focused on the polymorphism of the 48 bp VNTR in exon III in the *DRD4* gene, which has been demonstrated to associate with executive functions (e.g., Kegel & Bus, 2013). We aimed to test the hypothesis that *DRD4* long carriers demonstrate better performance in speech perception in IM conditions, but not during EM conditions. In Experiment 1, in a small sample that was not screened for neuropsychiatric disorders, we demonstrated that long carriers displayed better recognition performance than short homozygotes in noise conditions involved significant IM (2-talker babble) across a range of SNRs (−4 to −20 dB), while both groups performed comparably in EM conditions (pink noise). With a larger sample that was screened for neuropsychiatric disorders, Experiment 2 replicated and extended these findings. The long variant of *DRD4* gene was associated with better speech recognition in noise conditions involved significant IM (1- and 2-talker babble), but not in noise conditions that were primarily EM (8-talker babble and SSN). *DRD4* genetic variation explained about 3% of the difference observed in speech recognition in IM conditions. Taken together, in line with the hypothesis, our results

suggest a listening condition-specific (i.e. IM) advantage of *DRD4* long allele in speech perception.

As discussed in the introduction, unlike EM, IM produces central interferences, which are likely to interfere with listeners' ability to select the target and inhibiting/ignoring the influences from the interfering noises (Shinn-Cunningham, 2008). This means that release from IM arguably requires listeners to engage prefrontal function to a greater extent, when compared with EM. Thus, mechanistically, the listening condition-specific advantage of *DRD4* long carriers in IM conditions is possibly related to their advantage in executive function, specifically executive attention/working memory capacity (Gorlick, et al., 2014), which is related to the processes to maintain information in short-term memory in the face of interference (Conway, et al., 2005; Engle, Tuholski, Laughlin, & Conway, 1999). Indeed, our mediation analysis demonstrated that this listening condition-specific advantage in IM conditions was mediated by enhanced executive attention/working memory capacity as measured with OSPAN.

4.2 “Vulnerability” gene vs. plasticity gene, and *DRD4* gene

Recently, it has been suggested that *DRD4* behaves not as a “vulnerability” gene but as a “plasticity” gene (Belsky et al., 2009; Wells, Beevers, Knopik, & McGue, 2013). The *vulnerability* gene hypothesis conceptualizes the long variant of *DRD4* as a risk factor for psychiatric disorders such as ADHD (Faraone, Doyle, Mick, & Biederman, 2001), and the likelihood of that disorder will increase in the face of adverse environment (Belsky & Hartman, 2014). In contrast, the *plasticity* gene idea contends that, rather than being more susceptible to adverse environmental influences, long allele carriers show increased susceptibility to environmental influences in general (Belsky & Hartman, 2014). For example, the *DRD4* long allele has been associated with higher levels of inattention in the context of insensitive early maternal care, but also with lower levels of inattention in the context of highly sensitive maternal care (Berry, Deater-Deckard, McCartney, Wang, & Petrill, 2013). This enhanced general susceptibility to environmental stimuli in long allele carriers may be driven by the heightened attention to contextually relevant information (Wells, et al., 2013). Hence, long carriers relative to short homozygotes may exhibit enhanced attention to goal-relevant information even in the face of interference (i.e., enhanced executive attention/working memory capacity, Gorlick, et al., 2014), which may lead to advantages in tasks that require similar processes. Indeed, findings from the current study showed that *DRD4* long carriers demonstrated better speech perception performance in IM conditions, which places greater demand on executive attention/working memory capacity. Together, our results are consistent with the argument that *DRD4* behaves not as a “vulnerability” gene (Belsky, et al., 2009; Wells, et al., 2013), and the presence of long allele variant may confer advantages in situations that engage goal-directed attention (Gorlick, et al., 2014).

4.3 Energetic masking vs. informational masking, and executive function

The contributions of cognitive abilities to individual's capacity to perceive speech in adverse conditions have received considerable interest recently (Chandrasekaran, et al., in press; Mattys, Davis, Bradlow, & Scott, 2012; Rönnberg et al., 2013; Rönnberg, Rudner, Foo, &

Lunner, 2008; Rönnerberg, Rudner, Lunner, & Zekveld, 2010; Stenfelt & Rönnerberg, 2009). Working memory has been one of the foci, and it has been argued to be the most significant cognitive predictor of capacity to understand speech in noise (Akeroyd, 2008). In studies with normal-hearing population, higher working memory capacity predicts better recognition performance in noise conditions primarily involving IM such as 1-talker babble (Koelewijn, et al., 2012; Zekveld, et al., 2013), but not in conditions predominantly involving EM such as SSN (Besser, et al., 2013; Koelewijn, et al., 2012; Zekveld, et al., 2012; Zekveld, et al., 2013). The dissociation between IM and EM in cognitive processes are further supported by our results, where enhanced executive attention/working memory capacity was associated with better speech recognition in noise conditions that involved significant IM (1- and 2-talker babble), but not in noise conditions primarily involving EM (8-talker babble and SSN).

Unlike working memory capacity, short-term memory has been revealed to be a less reliable predictor of recognition performance in noise. Most existing studies using digit span task only investigated noise situations that primarily involved EM, such as 6-talker babble (Gordon-Salant, Yeni-Komshian, Fitzgibbons, Cohen, & Waldroup, 2013; Tamati, Gilbert, & Pisoni, 2013) and SSN (Gordon-Salant, et al., 2013; Humes et al., 1994; Kronenberger et al., 2013). Most of these studies did not find a link between digit span task (including backward and forward) and recognition performance, although one recent study from Tamati et al. (2013) showed that these two span tasks were associated with recognition performance in 6-talker babble conditions when using a high-variability sentence recognition test (Gilbert, et al., 2013).

With regard to inhibitory control, prior research did not find significant correlation between Stroop test and recognition performance in masking conditions ranging from primarily IM to primarily EM, including 2-talker babble (Desjardins & Doherty, 2013), 6-talker babble (Desjardins & Doherty, 2013; Tamati, et al., 2013), and SSN (Desjardins & Doherty, 2013). The current study replicated these findings. As suggested by Tamati et al. (2013), the null effects observed for Stroop test may be due to that the test does not challenge the listeners much, so there are not sufficient meaningful performance variations on this test for any relationship to emerge.

Taken together, previous findings and our current results suggest that IM requires a specific, highly-complex cognitive processing that may not be captured by simple short-term memory span and Stroop tests. Complex span tasks such as OSPAN capture the complexity by testing how participants hold on to information in the face of interference, which may be a better candidate to predict speech perception in IM conditions.

4.4 Energetic masking vs. informational masking, and SNR

Our results suggest that SNR exerts different effects on speech perception in these two masking conditions (Brungart, 2001; Freyman, et al., 1999). Specifically, the energetic masker overlaps with the target speech signals spectro-temporally at the periphery, resulting in degraded neural representation of the signals (Arbogast, et al., 2002; Brungart, 2001; Freyman, et al., 2004; Freyman, et al., 1999). As SNR decreases, these detrimental effects become more severe; hence, a rapid drop-off in performance was observed in this

experiment. In IM, however, even at quite adverse SNR, some amount of target speech information is still available to the listeners via “glimpsing”- the spectrotemporal regions where the maskers have a minimal impact on the target speech signals (Cooke, 2006). Thus, performance in IM is less affected by SNR.

4.5 Limitations of current study

It should be noted that other factors, co-varying with the genetic factor, may explain the individual differences in speech-in-noise performance. Hearing-related factors could be one explanatory variable. For example, although participants in the current study were classified as “normal hearing” based on pure-tone hearing threshold test, they may suffer from King-Kopetzky syndrome (e.g., Zhao & Stephens, 2007), which is characterized by difficulty in perceiving speech in noise but with clinically-normal pure-tone hearing threshold. However, we believe that the set of results (condition-specific effects) and a replication with a wide range of SNRs is reassuring regarding a contribution of *DRD4* variation on speech-in-noise performance. If hearing acuity-related factors play a role here, we may predict that the group differences would extend to EM conditions and be more substantial when the SNR is poorer. However, our results did not support this prediction. Hence, it is unlikely that audiometric thresholds could be a factor that explains the results. Nevertheless, future studies investigating auditory processing should test participants’ audiometric threshold more thoroughly.

Further, the association between the capacity to understand speech in IM conditions and *DRD4* should be interpreted with several limitations in mind. First, this association may be driven by another genetic variant in linkage disequilibrium with *DRD4* exon III VNTR. In addition, even though we analyzed the Caucasian sample, population stratification could still be considered as a possible explanation for the observed effects (Hutchison, Stallings, McGeary, & Bryan, 2004). Second, the samples of the two experiments in the current study departed from Hardy Weinberg Equilibrium (HWE). As such, results in this study should be interpreted with caution, because our sample may represent a non-random sample of the population. A number of situations may cause the departures from HWE such as genotyping error, nonrandom mating, and selected samples. It is unknown which factor may be responsible here. Genotyping error is unlikely, because we have high quality control checks in place in the lab, and we have genotyped for this same polymorphism for many other samples genotyped that have not differed from HWE. It is also unlikely that the departure from HWE resulted from selected samples, because the current sample was recruited from the community and was not selected based on any phenotype that would induce such results. Third, for a genetic association study, the replication sample in experiment 2 is still relatively small. Thus, a larger sample would be needed to further increase our confidence in the findings reported in this study.

4.6 Conclusion

Two experiments demonstrated and replicated the findings that individuals with the long variant in exon III of the *DRD4* gene exhibit a *listening condition-specific* advantage in speech perception under listening conditions that involve IM. This *listening condition-specific* advantage is mediated by enhanced working memory capacity in individuals with

the long allele variant. Despite acknowledged limitations, this foundational work provides important new insight into genetic influences on individual variability in the domain of speech perception in adverse conditions.

Acknowledgements

This work was supported by NIDA grant DA032457 to WTM. Research reported in this publication was also supported by the National Institute On Deafness And Other Communication Disorders of the National Institutes of Health under Award Number R01DC013315 (awarded to BC). This material is the result of work supported with resources and the use of facilities at the Providence Veterans Affairs Medical Center. The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government. We thank the Maddox Lab RAs for all data collection. We also thank Kristin J. Van Engen, Kirsten Smayda, Han-Gyol Yi, and Jasmine E. B. Phelps for their invaluable assistance in stimulus preparation, data management, and data analysis.

References

- Aben B, Stapert S, Blokland A. About the distinction between working memory and short-term memory. *Frontiers in Psychology*. 2012; 3 doi: Artn 301Doi 10.3389/Fpsyg.2012.00301.
- Akeroyd MA. Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*. 2008; 47(S2):S53–S71. [PubMed: 19012113]
- Altink ME, Rommelse NN, Slaats-Willemse DI, Vázquez AA, Franke B, Buschgens CJ, Oosterlaan J. The dopamine receptor D4 7-repeat allele influences neurocognitive functioning, but this effect is moderated by age and ADHD status: an exploratory study. *World Journal of Biological Psychiatry*. 2012; 13(4):293–305. [PubMed: 22111665]
- Alvarez JA, Emory E. Executive function and the frontal lobes: a meta-analytic review. *Neuropsychology review*. 2006; 16(1):17–42. [PubMed: 16794878]
- Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N. A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hearing research*. 2013; 300:18–32. [PubMed: 23541911]
- Arbogast TL, Mason CR, Kidd G. The effect of spatial separation on informational and energetic masking of speech. *Journal of the Acoustical Society of America*. 2002; 112(5):2086–2098. doi: Doi 10.1121/1.1510141. [PubMed: 12430820]
- Asghari V, Sanyal S, Buchwaldt S, Paterson A, Jovanovic V, Van Tol HH. Modulation of intracellular cyclic AMP levels by different human dopamine D4 receptor variants. *Journal of neurochemistry*. 1995; 65(3):1157–1165. [PubMed: 7643093]
- Audacity Developer Team. Audacity (Version 1.2. 6)[Computer software]. 2008. Available: audacity.sourceforge.net/download
- Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*. 2008; 59(4):390–412.
- Bamford J, Wilson I. Methodological considerations and practical aspects of the BKB sentence lists. *Speech-hearing tests and the spoken language of hearing-impaired children*. 1979:148–187.
- Barnes JJ, Dean AJ, Nandam LS, O'Connell RG, Bellgrove MA. The molecular genetics of executive function: role of monoamine system genes. *Biological Psychiatry*. 2011; 69(12):e127–e143. [PubMed: 21397212]
- Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*. 1986; 51(6):1173. [PubMed: 3806354]
- Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*. 2013; 68(3):255–278.
- Bates, D.; Maechler, M.; Bolker, B. lme4: Linear mixed-effects models using S4 classes. 2012.
- Bellgrove MA, Hawi Z, Kirley A, Gill M, Robertson IH. Dissecting the attention deficit hyperactivity disorder (ADHD) phenotype: Sustained attention, response variability and spatial attentional

- asymmetries in relation to dopamine transporter (DAT1) genotype. [Article]. *Neuropsychologia*. 2005; 43(13):1847–1857. doi: 10.1016/j.neuropsychologia.20.05.03.01. [PubMed: 16168728]
- Belsky J, Hartman S. Gene-environment interaction in evolutionary perspective: differential susceptibility to environmental influences. *World Psychiatry*. 2014; 13(1):87–89. [PubMed: 24497257]
- Belsky J, Jonassaint C, Pluess M, Stanton M, Brummett B, Williams R. Vulnerability genes or plasticity genes? *Molecular psychiatry*. 2009; 14(8):746–754. [PubMed: 19455150]
- Berry D, Deater-Deckard K, McCartney K, Wang Z, Petrill SA. Gene– environment interaction between dopamine receptor D4 7-repeat polymorphism and early maternal sensitivity predicts inattention trajectories across middle childhood. *Development and psychopathology*. 2013; 25(02): 291–306. [PubMed: 23627945]
- Besser J, Koelewijn T, Zekveld AA, Kramer SE, Festen JM. How linguistic closure and verbal working memory relate to speech recognition in noise—a review. *Trends in amplification*. 2013; 17(2):75–93. [PubMed: 23945955]
- Boersma, P.; Weenink, D. Praat: doing phonetics by computer (Version 5.1. 25)[Computer program]. 2010. Retrieved January 20, 2010
- Boonstra AM, Kooij J, Buitelaar JK, Oosterlaan J, Sergeant JA, Heister J, Franke B. An exploratory study of the relationship between four candidate genes and neurocognitive performance in adult ADHD. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2008; 147(3): 397–402.
- Bouchard TJ, Lykken DT, McGue M, Segal NL, Tellegen A. Sources of human psychological differences: The Minnesota study of twins reared apart. *Science*. 1990; 250(4978):223–228. [PubMed: 2218526]
- Bradlow AR, Alexander JA. Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *Journal of the Acoustical Society of America*. 2007; 121(4):2339–2349. doi: 10.1121/1.2642103. [PubMed: 17471746]
- Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America*. 2001; 109(3):1101–1109. doi: 10.1121/1.1345696. [PubMed: 11303924]
- Brungart DS, Chang PS, Simpson BD, Wang D. Multitalker speech perception with ideal time-frequency segregation: Effects of voice characteristics and number of talkers. *The Journal of the Acoustical Society of America*. 2009; 125(6):4006–4022. [PubMed: 19507982]
- Calandruccio L, Smiljanic R. New sentence recognition materials developed using a basic non-native English lexicon. *Journal of Speech, Language, and Hearing Research*. 2012; 55(5):1342–1355.
- Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N. Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron*. 2009; 64(3):311–319. [PubMed: 19914180]
- Chandrasekaran B, Van Engen K, Xie Z, Beevers CG, Maddox WT. Influence of depressive symptoms on speech perception in adverse listening conditions. *Cognition & Emotion*. in press.
- Collette F, Van der Linden M. Brain imaging of the central executive component of working memory. *Neuroscience & Biobehavioral Reviews*. 2002; 26(2):105–125. [PubMed: 11856556]
- Congdon E, Lesch KP, Canli T. Analysis of DRD4 and DAT polymorphisms and behavioral inhibition in healthy adults: implications for impulsivity. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 2008; 147(1):27–32.
- Conway AR, Kane MJ, Bunting MF, Hambrick DZ, Wilhelm O, Engle RW. Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*. 2005; 12(5): 769–786. [PubMed: 16523997]
- Cooke M. A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*. 2006; 119(3):1562–1573. [PubMed: 16583901]
- Cooke M, Lecumberri MLG, Barker J. The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *Journal of the Acoustical Society of America*. 2008; 123(1):414–427. doi: 10.1121/1.2804952. [PubMed: 18177170]
- Cools R, D'Esposito M. Inverted-U–Shaped Dopamine actions on human working memory and cognitive control. *Biological Psychiatry*. 2011; 69(12):e113–e125. [PubMed: 21531388]

- Cools R, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M. Working memory capacity predicts dopamine synthesis capacity in the human striatum. *The Journal of Neuroscience*. 2008; 28(5): 1208–1212. [PubMed: 18234898]
- Desjardins JL, Doherty KA. Age-related changes in listening effort for various types of masker noises. *Ear and hearing*. 2013; 34(3):261–272. [PubMed: 23095723]
- Ding Y-C, Chi H-C, Grady DL, Morishima A, Kidd JR, Kidd KK, Swanson JM. Evidence of positive selection acting at the human dopamine receptor D4 gene locus. *Proceedings of the National Academy of Sciences*. 2002; 99(1):309–314.
- Engels WR. Exact tests for Hardy–Weinberg proportions. *Genetics*. 2009; 183(4):1431–1441. [PubMed: 19797043]
- Engle RW, Tuholski SW, Laughlin JE, Conway AR. Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*. 1999; 128(3):309. [PubMed: 10513398]
- Faraco CC, Unsworth N, Langley J, Terry D, Li KM, Zhang DG, Miller LS. Complex span tasks and hippocampal recruitment during working memory. *Neuroimage*. 2011; 55(2):773–787. doi: DOI 10.1016/j.neuroimage.2010.12.033. [PubMed: 21182968]
- Faraone SV, Doyle AE, Mick E, Biederman J. Meta-analysis of the association between the 7-repeat allele of the dopamine D4 receptor gene and attention deficit hyperactivity disorder. *American Journal of Psychiatry*. 2001; 158(7):1052–1057. [PubMed: 11431226]
- Freyman RL, Balakrishnan U, Helfer KS. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America*. 2004; 115(5):2246–2256. doi: Doi 10.1121/1.689343. [PubMed: 15139635]
- Freyman RL, Helfer KS, McCall DD, Clifton RK. The role of perceived spatial separation in the unmasking of speech. *Journal of the Acoustical Society of America*. 1999; 106(6):3578–3588. doi: Doi 10.1121/1.428211. [PubMed: 10615698]
- Friedman NP, Miyake A, Young SE, DeFries JC, Corley RP, Hewitt JK. Individual differences in executive functions are almost entirely genetic in origin. *Journal of experimental psychology: General*. 2008; 137(2):201. [PubMed: 18473654]
- Gilbert JL, Tamati TN, Pisoni DB. Development, Reliability and Validity of PRESTO: A New High-Variability Sentence Recognition Test. *Journal of the American Academy of Audiology*. 2013; 24(1):26. [PubMed: 23231814]
- Gilsbach S, Neufang S, Scherag S, Vloet TD, Fink GR, Herpertz-Dahlmann B, Konrad K. Effects of the DRD4 genotype on neural networks associated with executive functions in children and adolescents. *Developmental cognitive neuroscience*. 2012; 2(4):417–427. [PubMed: 22727763]
- Gizer IR, Waldman ID. Double Dissociation Between Lab Measures of Inattention and Impulsivity and the Dopamine Transporter Gene (DAT1) and Dopamine D4 Receptor Gene (DRD4). *Journal of Abnormal Psychology*. 2012; 121(4):1011–1023. doi: Doi 10.1037/A0028225. [PubMed: 22564179]
- Gordon-Salant S, Yeni-Komshian GH, Fitzgibbons PJ, Cohen JI, Waldroup C. Recognition of accented and unaccented speech in different maskers by younger and older listeners. *The Journal of the Acoustical Society of America*. 2013; 134(1):618–627. [PubMed: 23862836]
- Gorlick, MA.; Worthy, DA.; Knopik, VS.; McGeary, JE.; Beevers, CG.; Maddox, WT. DRD4 Long Allele Carriers Show Heightened Attention to High-priority Items Relative to Low-priority Items. 2014.
- Humes LE, Watson BU, Christensen LA, Cokely CG, Halling DC, Lee L. Factors associated with individual differences in clinical measures of speech recognition among the elderly. *Journal of Speech, Language, and Hearing Research*. 1994; 37(2):465–474.
- Hutchinson KE, McGeary J, Smolen A, Bryan A, Swift RM. The DRD4 VNTR polymorphism moderates craving after alcohol consumption. *Health Psychology*. 2002; 21(2):139. [PubMed: 11950104]
- Hutchison KE, Stallings M, McGeary J, Bryan A. Population stratification in the candidate gene study: fatal threat or red herring? *Psychological Bulletin*. 2004; 130(1):66. [PubMed: 14717650]
- Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. *Nature genetics*. 2001; 29(3):306–309. [PubMed: 11600885]

- Kane MJ, Engle RW. The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*. 2002; 9(4):637–671. doi: Doi 10.3758/Bf03196323. [PubMed: 12613671]
- Kegel CA, Bus AG. Links between DRD4, executive attention, and alphabetic skills in a nonclinical sample. *Journal of Child Psychology and Psychiatry*. 2013; 54(3):305–312. [PubMed: 22897338]
- Kieling C, Roman T, Doyle AE, Hutz MH, Rohde LA. Association between DRD4 gene and performance of children with ADHD in a test of sustained attention. *Biological Psychiatry*. 2006; 60(10):1163–1165. doi: DOI 10.1016/j.biopsych.2006.04.027. [PubMed: 16781678]
- Koelewijn T, Zekveld AA, Festen JM, Rönnberg J, Kramer SE. Processing load induced by informational masking is related to linguistic abilities. *International journal of otolaryngology*. 2012; 2012
- Krämer UM, Rojo N, Schüle R, Cunillera T, Schöls L, Marco-Pallarés J, Münte TF. ADHD candidate gene (DRD4 exon III) affects inhibitory control in a healthy sample. *BMC neuroscience*. 2009; 10(1):150. [PubMed: 20021692]
- Kronenberger WG, Pisoni DB, Harris MS, Hoen HM, Xu H, Miyamoto RT. Profiles of verbal working memory growth predict speech and language development in children with cochlear implants. *Journal of Speech, Language, and Hearing Research*. 2013; 56(3):805–825.
- Landau SM, Lal R, O'Neil JP, Baker S, Jagust WJ. Striatal dopamine and working memory. *Cerebral Cortex*. 2009; 19(2):445–454. [PubMed: 18550595]
- Langley K, Marshall L, van den Bree M, Thomas H, Owen M, O'Donovan M, Thapar A. Association of the dopamine D4 receptor gene 7-repeat allele with neuropsychological test performance of children with ADHD. *American Journal of Psychiatry*. 2004; 161(1):133–138. [PubMed: 14702261]
- Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Harnett Sheehan K, Dunbar G. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry*. 1997; 12(5):224–231.
- Li S-C, Passow S, Nietfeld W, Schröder J, Bertram L, Heekeren HR, Lindenberger U. Dopamine modulates attentional control of auditory perception: DARPP-32 (PPP1R1B) genotype effects on behavior and cortical evoked potentials. *Neuropsychologia*. 2013; 51(8):1649–1661. [PubMed: 23639477]
- Lichter JB, Barr CL, Kennedy JL, Van Tol HH, Kidd KK, Livak KJ. A hypervariable segment in the human dopamine receptor D4 (DRD4) gene. *Human Molecular Genetics*. 1993; 2(6):767–773. [PubMed: 8353495]
- Loo SK, Rich EC, Ishii J, McGough J, McCracken J, Nelson S, Smalley SL. Cognitive functioning in affected sibling pairs with ADHD: familial clustering and dopamine genes. *Journal of Child Psychology and Psychiatry*. 2008; 49(9):950–957. [PubMed: 18665883]
- Marian V, Blumenfeld HK, Kaushanskaya M. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*. 2007; 50(4):940–967.
- Mattys SL, Davis MH, Bradlow AR, Scott SK. Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*. 2012; 27(7-8):953–978.
- Oak JN, Oldenhof J, Van Tol HHM. The dopamine D-4 receptor: one decade of research. *European Journal of Pharmacology*. 2000; 405(1-3):303–327. doi: Doi 10.1016/S0014-2999(00)00562-8. [PubMed: 11033337]
- Parbery-Clark A, Strait D, Kraus N. Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*. 2011; 49(12):3338–3345. [PubMed: 21864552]
- Rönnberg J, Lunner T, Zekveld A, Sörqvist P, Danielsson H, Lyxell B, Pichora-Fuller MK. The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in systems neuroscience*. 2013; 7
- Rönnberg J, Rudner M, Foo C, Lunner T. Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*. 2008; 47(S2):S99–S105. [PubMed: 19012117]

- Rönnerberg J, Rudner M, Lunner T, Zekveld AA. When cognition kicks in: Working memory and speech understanding in noise. *Noise and Health*. 2010; 12(49):263. [PubMed: 20871181]
- Seamans JK, Yang CR. The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Progress in neurobiology*. 2004; 74(1):1–58. [PubMed: 15381316]
- Sheehan D, Lecrubier Y, Harnett Sheehan K, Janavs J, Weiller E, Keskiner A, Dunbar G. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry*. 1997; 12(5):232–241.
- Shinn-Cunningham BG. Object-based auditory and visual attention. *Trends in Cognitive Sciences*. 2008; 12(5):182–186. doi: DOI 10.1016/j.tics.2008.02.003. [PubMed: 18396091]
- Simpson SA, Cooke M. Consonant identification in N-talker babble is a nonmonotonic function of N. *J. Acoust. Soc. Am*. 2005; 118(5):2775–2778. [PubMed: 16334654]
- Song JH, Skoe E, Banai K, Kraus N. Perception of Speech in Noise: Neural Correlates. *Journal of Cognitive Neuroscience*. 2011; 23(9):2268–2279. [PubMed: 20681749]
- Stenfelt S, Rönnerberg J. The Signal-Cognition interface: Interactions between degraded auditory signals and cognitive processes. *Scandinavian journal of psychology*. 2009; 50(5):385–393. [PubMed: 19778386]
- Stroop JR. Studies of interference in serial verbal reactions. *Journal of experimental psychology*. 1935; 18(6):643.
- Sussman ES, Bregman AS, Lee W-W. Effects of task-switching on neural representations of ambiguous sound input. *Neuropsychologia*. 2014
- Swanson J, Oosterlaan J, Murias M, Schuck S, Flodman P, Spence MA, Smith M. Attention deficit/hyperactivity disorder children with a 7-repeat allele of the dopamine receptor D4 gene have extreme behavior but normal performance on critical neuropsychological tests of attention. *Proceedings of the National Academy of Sciences*. 2000; 97(9):4754–4759.
- Takahashi H, Kato M, Takano H, Arakawa R, Okumura M, Otsuka T, Ito H. Differential contributions of prefrontal and hippocampal dopamine D1 and D2 receptors in human cognitive functions. *The Journal of Neuroscience*. 2008; 28(46):12032–12038. [PubMed: 19005068]
- Tamati TN, Gilbert JL, Pisoni DB. Some factors underlying individual differences in speech recognition on PRESTO: A first report. *Journal of the American Academy of Audiology*. 2013; 24(7):616. [PubMed: 24047949]
- Unsworth N, Engle RW. Working memory capacity and fluid abilities: Examining the correlation between Operation Span and Raven. *Intelligence*. 2005; 33(1):67–81.
- Van Engen KJ. Speech-in-speech recognition: A training study. *Language and Cognitive Processes*. 2012; 27(7-8):1089–1107. doi: Doi 10.1080/01690965.2012.654644.
- Van Engen KJ, Baese-Berk M, Baker RE, Choi A, Kim M, Bradlow AR. The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and speech*. 2010; 53(4):510–540. [PubMed: 21313992]
- Van Engen KJ, Phelps JE, Smiljanic R, Chandrasekaran B. Enhancing speech intelligibility: interactions among context, modality, speech style, and masker. *Journal of Speech, Language, and Hearing Research*. 2014
- Van Tol, HH.; Wu, CM.; Guan, H-C.; Ohara, K.; Bunzow, JR.; Civelli, O.; Jovanovic, V. Multiple dopamine D4 receptor variants in the human population. 1992.
- Vijayraghavan S, Wang M, Birnbaum SG, Williams GV, Arnsten AF. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nature neuroscience*. 2007; 10(3):376–384.
- Wechsler, D. WAIS-III, Wechsler Adult Intelligence Scale: Administration and Scoring Manual. Psychological Corporation; 1997.
- Wells TT, Beevers CG, Knopik VS, McGeary JE. Dopamine D4 receptor gene variation is associated with context-dependent attention for emotion stimuli. *The International Journal of Neuropsychopharmacology*. 2013; 16(03):525–534. [PubMed: 22607734]
- Wightman FL, Kistler DJ, O'Bryan A. Individual differences and age effects in a dichotic informational masking paradigm. *Journal of the Acoustical Society of America*. 2010; 128(1):270–279. doi: Doi 10.1121/1.3436536. [PubMed: 20649222]

- Wilson RH, McArdle RA, Smith SL. An evaluation of the BKB-SIN, HINT, QuickSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *Journal of Speech, Language, and Hearing Research*. 2007; 50(4):844–856.
- Zekveld AA, Rudner M, Johnsrude IS, Heslenfeld DJ, Rönnberg J. Behavioral and fMRI evidence that cognitive ability modulates the effect of semantic context on speech intelligibility. *Brain and language*. 2012; 122(2):103–113. [PubMed: 22728131]
- Zekveld AA, Rudner M, Johnsrude IS, Rönnberg J. The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *The Journal of the Acoustical Society of America*. 2013; 134(3):2225–2234. [PubMed: 23967952]
- Zhao F, Stephens D. A critical review of King-Kopetzky syndrome: Hearing difficulties, but normal hearing? *Audiological Medicine*. 2007; 5(2):119–124.

Fig. 1.

Boxplot of proportion of correctly identified keywords from -4 to -20 dB in *DRD4* short homozygotes (dark bar) and long carriers (light bar) in two-talker babble (A) and pink noise conditions (B). The horizontal line shows the median value, the boxes shows the quartiles, the whiskers represent 1.5 times the interquartile range, and the black dots depict outliers. Outliers are defined as cases with values between 1.5 and 3 times the interquartile range from the upper or lower edge of the box.

Fig. 2.

Boxplot of proportion of correctly identified keywords in the *DRD4* short homozygotes (dark bar) and long carriers (light bar) in informational masking and energetic masking conditions. The horizontal line shows the median value, the boxes shows the quartiles, the whiskers represent 1.5 times the interquartile range.

Table 1

Demographics of the sample for analysis in Experiment 1

	<i>DRD4</i> long carriers	<i>DRD4</i> short homozygotes
Age	18.8(1.03)	19.6(4.68)
Years of Education	12.70(0.95)	12.78 (0.89)
Gender	7 female, 3 male	27 female, 13 male
Ethnicity		
Hispanic	3	6
Non-Hispanic	7	33
Decline to State	0	1

Note: Standard deviations are listed in parentheses.

Table 2

Results of the linear mixed effects logistic regression on the intelligibility data in *DRD4* long carriers and short homozygotes in two-talker babble and pink noise conditions in Experiment 1

Fixed effects:	β	<i>SE</i>	<i>Z</i> value	<i>p</i>
(Intercept)	−0.21	−0.21	−1.68	.092
Condition (noise_Pink noise)	−1.01	0.07	−13.67	< .001
Genotype (<i>DRD4</i> _Long carriers)	0.64	0.17	3.80	< .001
SNR	0.22	0.008	29.15	< .001
Condition : Genotype (noise_Pink noise: <i>DRD4</i> _Long carriers)	−0.51	0.15	−3.41	.001
Condition : SNR (noise_Pink noise:SNR)	0.19	0.01	13.55	< .001

Table 3

Demographics of the sample for analysis in Experiment 2

	<i>DRD4</i> long carriers	<i>DRD4</i> short homozygotes
Age	25.6(4.24)	25.83(4.47)
Years of Education	16.23(2.74)	15.69 (2.55)
Gender	12 female, 8 male	46 female, 38 male
Ethnicity		
Hispanic	2	12
Non-Hispanic	17	71
Decline to State	1	1

Note: Standard deviations are listed in parentheses.

Table 4

Results of the linear mixed effects logistic regression on the intelligibility data across informational masking (IM) and energetic masking (EM) conditions in *DRD4* long carrier and short homozygotes.

Fixed effects:	β	<i>SE</i>	<i>Z</i> value	<i>p</i>
(Intercept)	0.32	0.07	4.33	< .001
Condition (Noise_EM)	0.59	0.03	17.37	< .001
Genotype (<i>DRD4</i> _Long carriers)	0.54	0.14	3.77	< .001
Condition : Genotype (Noise_EM: <i>DRD4</i> _Long carriers)	−0.51	0.08	−6.46	< .001

Table 5

Testing OSPAN as mediator between *DRD4* genotype and speech recognition performance in informational masking (IM) conditions

Testing steps in mediation model	β	<i>SE</i>	<i>t</i> value	<i>p</i>	Adjusted <i>R</i> ²
step1					
Outcome: Performance in IM					
Predictor: <i>DRD4</i> genotype	0.12	0.05	2.56	.011	.026
step2					
Outcome: OSPAN					
Predictor: <i>DRD4</i> genotype	9.87	4.06	2.43	.017	.050
step3					
Outcome: Performance in IM					
Predictor: OSPAN scores	0.004	0.001	3.93	< .001	.065
step4					
Outcome: Performance in IM					
Predictor: <i>DRD4</i> genotype	0.14	0.17	0.84	.404	
Predictor: OSPAN	0.004	0.001	3.30	.001	
Predictor: OSPAN : <i>DRD4</i> genotype	−0.001	0.003	−0.37	.714	
Total					.070
step5					
Outcome: performance in IM					
Predictor: <i>DRD4</i> genotype via OSPAN			2.07	.039	