Chapter 10
Dyslexia: A Failure in Attaining Expert-Level Reading Due to Poor Formation of Auditory Predictions

Sagi Jaffe-Dax, Luba Daikhin, and Merav Ahissar

Abstract In this chapter, we present evidence that alters the way Dyslexia is typically viewed and assessed. Based on accumulating findings obtained from behavioral assessments, computational modeling, imaging and ERP studies, we propose that Dyslexia results from a failure in acquiring a specific (reading and linguistic) skill that relies heavily on familiarity with stimuli distributions characterized by temporal regularities in a specific time window. Dyslexia is naturally associated with language related impairments, since learning temporal regularities is crucial for acquiring linguistic skills, but not confined to them. Studying Dyslexics' basic auditory processing from this perspective reveals specific and robust deficits in benefiting from simple temporal consistencies, which are associated with a reduced ability to accumulate stimuli statistics across time windows of >2–3 s. Importantly, similar impairments are demonstrated in the visual modality, supporting the cross-modal nature of the core deficit. Collectively, our findings show that Dyslexics fail to achieve expert level performance in variety of tasks, including reading, due to deficient accumulation of summary statistics, which impedes the formation of reliable predictions, which in turn facilitate switching performance to rely on efficient processing strategies.

Keywords Anchoring deficit hypothesis of dyslexia · Working memory · Perception · fMRI · EEG · Statistical learning

S. Jaffe-Dax (✉)
Princeton University, Princeton, NJ, USA
e-mail: jaffedax@princeton.edu

L. Daikhin
Department of Psychology, The Hebrew University, Jerusalem, Israel
e-mail: luba.daikhin@mail.huji.ac.il

M. Ahissar
Department of Psychology and Edmond and Lily Safra Center for Brain Sciences, The Hebrew University, Jerusalem, Israel
e-mail: merav.ahissar@mail.huji.ac.il

10.1 A General Theory of Skill Acquisition and Cognitive Disabilities

We propose a general, principled theory of skill acquisition, which integrates a theory of learning (Reverse Hierarchy Theory, RHT – Ahissar & Hochstein, 1997; 2004; Rokem & Ahissar, 2009), with the 2-system theory of modes of cognition (Kahneman 2011). RHT proposes a separation between bottom-up (local to global) processing, and top-down (global to local) perception, and perceptual learning. It asserts that as a default our attention system and, in accordance, our perception are based on high-level, object- and scene-oriented representations, which have an ecological reality. This is useful for every day purposes, but when learning requires fine discriminations (e.g., small bars, similar letters), a backward search for allocating informative neural populations, that input the higher level, is activated. When such populations are allocated, they can gradually replace higher-level populations as part of the overall scheme of performance, with better resolution. This process leads to "pushing down" initially high-level roles, with practice, but only when this practice is successful in detecting lower-level populations, i.e. detect and integrate regularities. The dual system approach to cognition proposes two modes of cognition: fast, automatic and effortless, but prone to "perceptual biases" (system 1) versus slow, serial and effortful (system 2). We propose that these are two extremes along a hierarchy of modes of performance that characterize the gradual skill-acquisition process from the novice (system 2) to the expert (system 1). System 2 is embodied by the working-memory fronto-parietal system, which is consistently activated in novel and challenging situations (Duncan & Owen 2000), and has been termed the "multiple-demand system" (Duncan 2010). This system was associated with general intelligence (e.g. the amount of space taken by these areas is correlated with reasoning skills; Woolgar et al. 2010).

The novice has to "set the stage" to perform a task, relying mainly on the domain-general high-level fronto-parietal system, which maps new tasks to neural implementations. Successful practice proceeds along a hierarchy of processing in a reverse direction, where gradually lower-level areas encode sub-parts of the trained tasks. Crucially, the gradual reliance on lower levels depends on allocating neuronal populations that reliably encode task-related regularities, which is the neural correlate of regularity detection. Thus, the expert does not activate the same algorithms faster, but replaces them with low-level encoded schemes.

Practice does not always make better. One has to detect the repeated regularities. We propose that specific developmental disabilities result from a failure to automatically detect and use sensory or sensory-motor regularities under conditions that allow such detection for the general population. What characterizes developmental disabilities is adequate general reasoning skills, allowing understanding of novel tasks, but reduced ability to attain fast and effortless performance based on retrieved schemes, in spite of intensive practice. We propose that this conceptual account applies to a variety of developmental disabilities, though here we shall focus on the case of reading disability. This focus allows us to study the computational and

neural mechanisms that yield these complex phenomena in an unprecedented depth, leading to conceptual shift in our understanding of disabilities as part of a principled theory of learning and inference.

Dyslexia is a pervasive difficulty in attaining expert-level reading, in spite of adequate reasoning skills and adequate practice and guidance (education) opportunities (World Health Organization 2008). Within our proposed framework we view dyslexia as follows: Dyslexics' 'system 2' is adequate. However, they have difficulties in delegating reading sub-tasks to gradually lower-level areas, due to inefficiency in detecting sound regularities. These difficulties impact the processing of both simple and complex sounds, and impede the acquisition of expert level reading due to inefficiency in representing linguistic regularities. In this chapter, we describe our series of observations and modeling supporting this hypothesis.

10.2 The Anchoring Hypothesis of Dyslexia

In the early 2000s, two open questions were heatedly debated in the field of dyslexia – (1) Are dyslexics' deficits specific to speech sounds? (2) Is the deficit representational or does it affect only access to otherwise adequate representations? In an early study we performed, we found that a large fraction of adult poor readers also perform poorly in a broad set of simple auditory discrimination, and that the degree of their deficit is correlated with their reading scores (Ahissar, Protopapas, Reid, & Merzenich 2000; Amitay, Ahissar, & Nelken 2002), suggesting that dyslexics' deficits are not specific to speech sounds.

Trying to better decipher the bottleneck underlying dyslexics' poor performance in simple discrimination tasks, we administered two protocols of 2-tone frequency discrimination asking which tone is higher (Ahissar, Lubin, Putter-Katz, & Banai 2006). In one protocol there was a fixed reference frequency in each trial whereas in the other, there was no such consistency. Listeners' discrimination thresholds were substantially lower when there was a reference, even though they were unaware of its presence. Our dyslexic participants did not show this benefit (Fig. 10.1). In a subsequent study, we administered the same two protocols to dyslexics and ADHD participants (Oganian & Ahissar 2012). We replicated the effect of impaired benefit from stimulus repetition among dyslexics, whereas individuals with ADHD who were good readers did not show this effect. All their thresholds were similar to controls'.

To test for a similar deficit underlying speech perception, we measured speech perception in noise under two related conditions: with a large set of words, and with a small set, hence many word repetitions. Dyslexics' deficits were only found in the small set, with repetitions. Importantly, their pattern of errors showed similar sensory sensitivity as controls (e.g., "/tarul/" instead of "/barul/"), but in contrast to controls, it was not restricted to the trained set of words, suggesting poor usage of word repetitions. We conclude that dyslexics' deficit in implicit learning of distributional statistics also results in the formation of impoverished categorical rep-
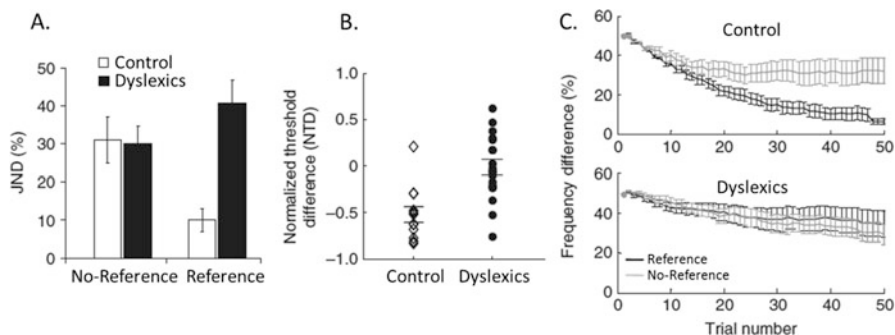
Fig. 10.1 Frequency discrimination thresholds with and without stimulus repetition across trials among dyslexics and control individuals. The effects of discrimination condition (No-reference versus Reference) differed significantly between the groups. (**a**) Average thresholds show that dyslexics had significantly higher JNDs in the Reference condition. (**b**) Single-subject data of the normalized difference in threshold between the Reference and the No-reference conditions. Filled circles: dyslexics; open diamonds: control. (**c**) Adaptive assessment protocol for control individuals (top) and dyslexics (bottom) in the two procedures shows a gradual effect of using a consistent reference for control individuals but not for dyslexics. Error bars denote SEM. (Adapted from Ahissar et al. 2006, with permission)

resentations. One manifestation of this deficit is poorer performance in demanding speech discrimination tasks that rely on rich categorical representations (Banai & Ahissar 2017).

10.3   Evidence for Attaining Expert-Level Characteristics When Regularities Can Be Easily Detected

Even though context is known to crucially affect perception and learning, its specific impact had not been studied. We therefore conducted two studies to better understand the process of regularity detection, and its impact on the implicit change of underlying strategy.

The first study (Nahum, Daikhin, Lubin, Cohen, & Ahissar 2010) concerned two main questions: (1) What kind of regularities can be utilized quickly, within a single training session? (2) Is the network underlying task performance modified when regularities are detected and used to improve performance? We designed four protocols of frequency discrimination with simple regularities – Reference 1st, in which the first tone in each trial had a fixed, reference frequency (1000 Hz) and the second was higher or lower; Reference 2nd, where the second tone was fixed; Implicit Reference, where a fixed reference tone was presented five times at the beginning of the session and after that, a single tone, which could be higher or lower, was presented on every trial (evenly distributed around this reference); and Reference interleaved, where odd trials were of type Reference 1st and even

trials were of type Reference 2nd. We also applied a No-reference protocol, in which no regularities were included. Each protocol was measured in a different group of control participants, to avoid effects of more than one session training. Figure 10.2a shows the dynamics of behavior and the thresholds obtained under each of the protocols. For the Reference-interleaved protocol, the thresholds of Reference 1st trials (odd) and Reference 2nd trials (even) were obtained and presented separately. The benefit of the repeated reference tone at a fixed temporal position was the greatest. Indeed, the protocol that yielded lowest thresholds was Reference 1st, showing a steeper slope, indicating faster learning of the repeated structure. Reference 2nd protocol yielded low thresholds when presented with no competitive options, but was more vulnerable to interference, as indicated by its poor performance under the interleaved protocol. These results indicate that utilization of the repeated reference is fast (evident already by the 10th trial) and highly beneficial. Yet, there are more and less preferred conditions, which produce different degrees of benefit (i.e., higher sensitivity to regularity at the onset of a trial, i.e., event).

The similarity between the thresholds obtained under the Implicit Reference protocol and other protocols with a reference presented explicitly, questions whether the on-line comparison of the externally presented stimuli is, in fact, the strategy that is used for solving the task, since under this protocol only one tone (non-reference) is presented in each trial. To test this question we measured the performance under the Reference 1st and Reference 2nd protocols while recording Event Related Potentials (ERPs). Specifically we asked whether the ERP component, which denotes perceptual characterization, is produced at a different timing in these two types of trials. We reasoned that if listeners detect the regularity of the structure, they can compare the non-reference (target) tone to the internal reference. But for that they should detect which is the informative, non-reference tone, i.e. the structure of the protocol. If this is indeed the case, we can detect it by tracking a change in the temporal position of the P3 component. This ERP component is produced when a task-related categorization is made. We thus asked whether a clear P3 would be formed after the first tone in the trial in Reference 2nd protocol. As shown in Fig. 10.2b, this is indeed what we found! Under Reference 1st protocol, as expected, the P3 component was elicited after the second tone, since categorization could not have been made before this tone was presented. However, under the Reference 2nd protocol, where the first tone was the informative tone and the second was the repeated reference, P3 was formed ∼300 ms after the first tone, even before the 2nd tone was presented, although participants' introspective was of on-line comparison, and their button press followed the 2nd tone. These results further show that when successfully detected, cross-trial regularity leads to a strategic shift in the operations underlying performance.

What is the role of cross-trial regularities in long-term learning? How do the specifically trained information structures (Reference 1st: repeated → new; Reference 2nd: new → repeated) affect long term learning? Does learning generalize with practice or, is it specific to the trained protocol? To test this we trained two groups of participants (Cohen, Daikhin, & Ahissar 2013). They were trained with each
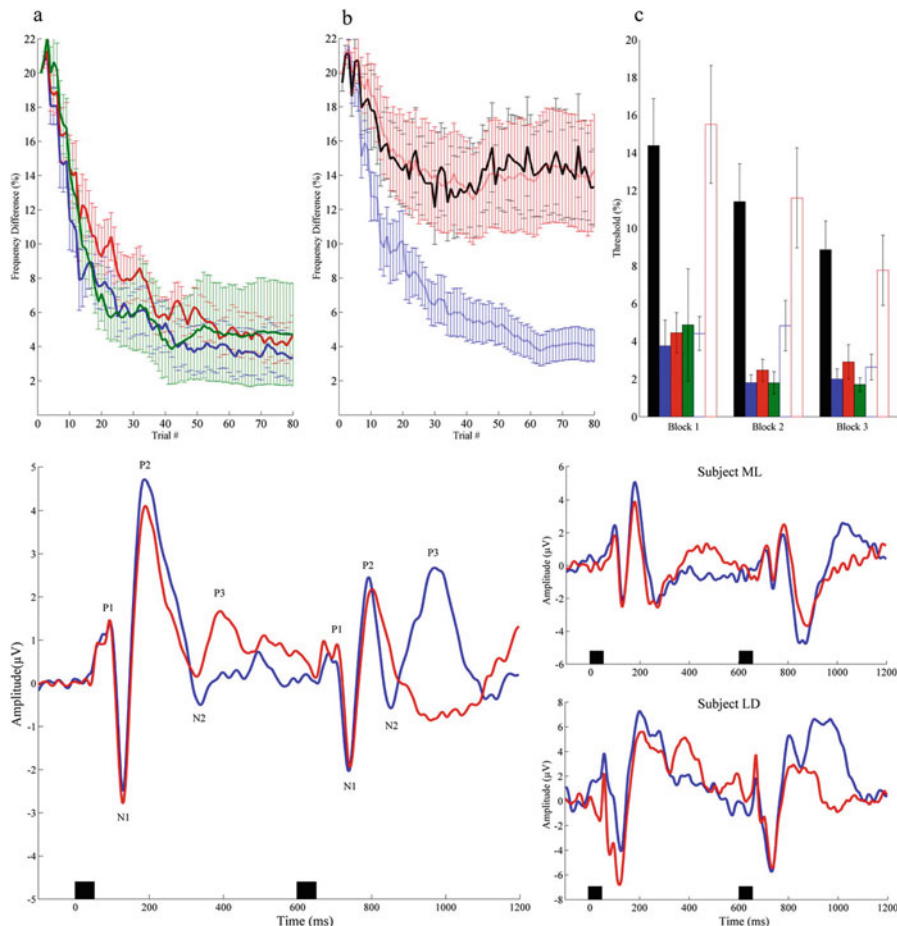
Fig. 10.2 (**a**) Behavioral performance under four different adaptive protocols: Left – frequency difference between tones as a function of trial number in the first assessment block under three protocols: reference 1st (blue), reference 2nd (red), and implicit reference (green). Middle – frequency difference along the first assessment for two additional protocols: No-reference (black, no repeated reference) and reference interleaved (reference 1st on odd trials (blue) and reference 2nd on even trials (red)) measured with the same adaptive paradigm. Right – thresholds (frequency difference) in three consecutive blocks performed with each of these five protocols. Although all thresholds showed improvement, their ranks were retained across assessments. Cross-subject averages and SEMs are shown. (**b**) ERP measured while participants performed the two-tone frequency discrimination task under reference 1st protocol (blue) and under reference 2nd protocol (red). The temporal location of the tones in a trial is marked by the black rectangles at the bottom of the plot. The relevant components are marked on the averaged waveforms. A clear P3 component can be seen after the second tone in the reference 1st protocol and after the first (non-reference) tone in the reference 2nd protocol. (Adapted from Nahum et al. 2010, with permission)
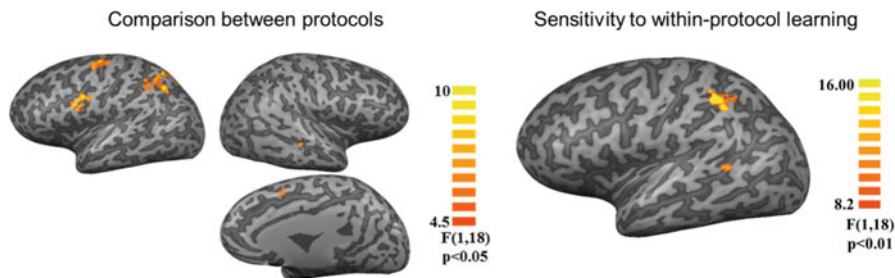
Fig. 10.3 Voxel-wise repeated-measures ANOVA (of beta values). Left – main effect of protocol; Right – main effect of within-protocol block (learning). (Adapted from Daikhin & Ahissar 2015, with permission)

of the two reference containing protocols – Reference 1st and Reference 2nd. We found that cross-session learning was largely specific to the structure of information in the trained protocol. Namely, learning did not transfer between protocols that shared stimuli, task, timing, but did not share the temporal structure of within trial regularity.

Nahum et al. (2010) study found that cross-trial regularities resulted in modification of the strategy used for solving the task. We hypothesized that detecting regularities would be associated with an increased role of auditory areas compared to the multiple-demand circuitry (i.e., task successfully delegated backward). To test this hypothesis, we conducted an fMRI study (Daikhin & Ahissar 2015), and compared the pattern of brain activation under Reference 1st (with regularity) and No-reference (no regularity) protocols, presented in an interleaved manner, so that participants were unaware of the protocol switch (every 36 trials). Figure 10.3 (left) shows the contrast between the patterns of activation under the two protocols. The cortical areas involved are associated with explicit working memory system and are part of the multiple-demand network (Duncan 2010). In each of these areas, the activation was higher under the No-reference protocol, in line with the hypothesis that utilizing the regularities leads to reduced load on the working memory system. Figure 10.3 (right) shows brain regions which were sensitive to within-protocol improvement across trials. Behaviorally, quick cross-trial improvement was found only for Reference 1st protocol. BOLD response changes were also found only under this protocol. Learning contrast was significant in two regions in the left hemisphere – an intra-parietal region, associated with controlling retention of information (Baldo & Dronkers 2006; Koelsch et al. 2009; Magen, Emmanouil, McMains, Kastner, & Treisman 2009; Sreenivasan, Curtis, & D'Esposito 2014), and a posterior superior temporal region, associated with regularity detection in simple and complex auditory stimuli (Binder et al. 2000; Davis & Johnsrude 2003; Friederici, Makuuchi, & Bahlmann 2009; Obleser & Kotz 2010). We interpret this modification as reflecting partial delegation of task performance to more posterior networks, which store effective sound regularities.

10.4 Integration of Previous Trials' Statistics in the More General Case: Contraction Bias

In the No-reference protocol listeners do not have specific item repetitions that can be used as anchors. However, even here listeners implicitly compute priors – representations of history-based knowledge, which substantially affect their performance. A related and amply documented phenomenon is "contraction bias", which is described as follows: when the magnitude of the two stimuli is small or low with respect to the mean of the previous stimuli in the experiment, participants tend to respond that the 2nd stimulus is smaller or lower, whereas when the magnitude of both stimuli is large or high they tend to respond that the 2nd stimulus is larger or higher (Hollingworth 1910; Preuschhof, Schubert, Villringer, & Heekeren 2010; Woodrow 1933). We have shown that this "contraction bias", can be understood within the Bayesian framework. Namely, participants form an integrated representation of the recently presented stimulus with the mean (prior) of previous trials. This integration is particularly helpful when responses are noisy and priors provide reliable predictions. Thus, one would expect that for noisier responses the weight of the priors in the integrated representation would be larger (Ashourian & Loewenstein 2011). The level of noise in the representation of the 1st stimulus is larger than the level of noise in the representation of the 2nd stimulus because of the additional noise associated with the encoding, and maintenance of the 1st stimulus in memory during the inter-stimulus interval of sequential presentation tasks (Bull & Cuddy 1972; Wickelgren 1969). Therefore, the integrated representation of the 1st tone in the trial is expected to be more biased (contracted) towards the calculated mean. Consequently, participants' responses are biased towards overestimating the 1st stimulus when it is small and underestimating it when it is large with respect to the prior.

The expected effect of the contraction bias on performance depends on the relative position of the first and second stimulus with respect to the distribution of stimuli in the experiment. Thus, there will be stimulus pairs that will gain from this bias (*Bias+*) and those that will lose (*Bias−*). *Bias+* trials are trials in which contraction of the first tone towards the mean frequency increases the difference between the representations of the two tones in the trial (Fig. 10.4, yellow zones). *Bias−* trials are those where such contraction decreases the perceived difference between the two stimuli, and hence hampers performance (gray zones in Fig. 10.4). In these trials, contracting the first tone towards the mean frequency decreases its perceived difference from the second tone and performance is thus expected to be impaired by this contraction. *Bias0* trials (white zones in Fig. 10.4) are trials in which the first and the second tones flank the mean frequency.
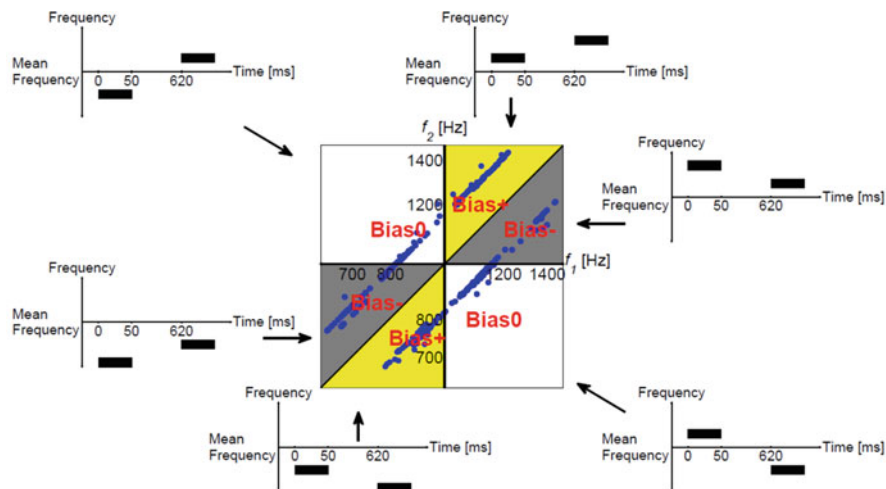
Fig. 10.4 Distribution of trials in the 2-tone discrimination task, and its impact on the perceived inter-pair frequency difference. Middle panel: Trial distributions presented by the frequency of the first and second tones [f_1, f_2] in each trial. Each dot denotes the f_1 and f_2 of a trial. The diagonal denotes $f_1 = f_2$. Equal distance lines from the diagonal denote trials with different frequencies but fixed within trial frequency difference, as plotted here. Surrounding schematic plots illustrate contraction bias. In *Bias+* trials the first tone is closer to the mean. Hence its contraction to the mean increases the perceived frequency difference between the two tones. In *Bias−* trials the first tone is farther from the mean, and contraction of the first tone decreases the perceived difference. In *Bias0* trials the two tones flank the mean

10.5   The Magnitude of Contraction Bias Is Smaller in Dyslexics Than in Controls

Raviv, Ahissar, and Loewenstein (2012) measured the magnitude of the contraction bias (the difference in success rate between *Bias+* and *Bias−* trials, as illustrated in Fig. 10.4) in the general population under the No-reference protocol and found a substantial effect. Jaffe-Dax, Raviv, Jacoby, Loewenstein, and Ahissar (2015) used a roughly fixed frequency difference (blue dots in Fig. 10.4), chosen as the difference that yields ~80% correct performance in good readers (as measured by Nahum et al. 2010). Though the difficulty of each trial was the same in terms of intra-trial frequency difference, success rate varied substantially across trials, in a manner that could be largely explained by the contraction bias (Fig. 10.5a).

We examined whether dyslexics' context effects were reduced by measuring the magnitude of their contraction bias compared with controls (Jaffe-Dax et al. 2015). Overall, dyslexics performed worse than controls. However, they showed a smaller context effect; i.e., a smaller difference in performance between *Bias+* and *Bias−* trials (Fig. 10.5b). Namely, in spite of their overall noisier representations, they under-weighted previous trials' statistics. Importantly, in some behavioral situations, where priors impair performance, dyslexics' performance gains from this
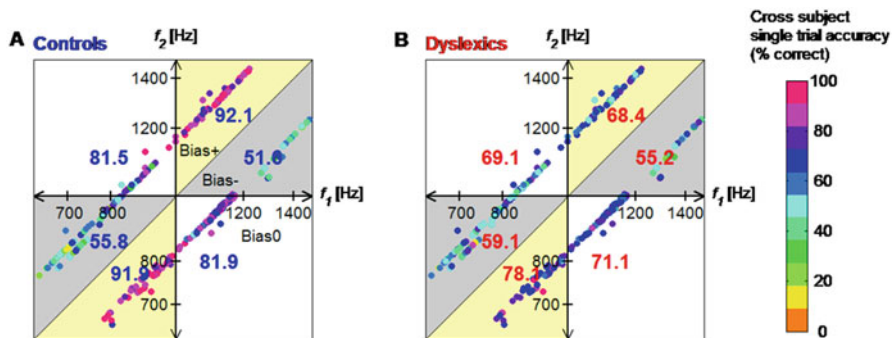
Fig. 10.5 Contraction bias (difference in performance between *Bias+* and *Bias−* zones) is larger in controls (left, average % correct in each zone in blue) than in dyslexics (right, average % correct in each zone in red). The color of each dot denotes the cross-subject average performance for that pair of stimuli. All participants were tested with the same stimuli set. Note the large color difference between dots in the *Bias+* and *Bias−* zones in the left plot (controls) versus the small difference in the right plot (dyslexics). Notations are the same as in Fig. 10.4. (Adapted from Jaffe-Dax et al. 2015, with permission)

implicit under-weighting. Thus, in the *Bias−* regions controls performed at chance level whereas dyslexics' performance was significantly above chance. Importantly, when time intervals between trials were manipulated, allowing the assessment of the dynamics of both behavioral and neural (adaptation) consequences, we found that implicit memory decays faster in dyslexics (Jaffe-Dax, Frenkel, & Ahissar 2017). Thus, dyslexics' retention is impaired with time intervals larger than ∼5 s.

10.6 Dyslexics' Implicit Memory Trace Is Less Sensitive to Stimulus' Statistics

We hypothesized that if dyslexics' reduced weighting of previous trials stems from an impaired formation of an integrated representation, reduced sensitivity to stimulus statistics may be apparent even before the second tone is presented. To test this hypothesis, we measured event-related potentials (ERPs). We focused on the dynamics and magnitude of the P2 component, which is an automatic response evoked by the auditory cortex (Mayhew, Dirckx, Niazy, Iannetti, & Wise 2010; Sheehan, McArthur, & Bishop 2005). Previous studies, utilizing both oddball (measuring mismatch negativity (MMN); Baldeweg 2007; Haenschel, Vernon, Dwivedi, Gruzelier, & Baldeweg 2005; Tong, Melara, & Rao 2009) and discrimination paradigms (Ross & Tremblay 2009; Tremblay, Inoue, McClannahan, & Ross 2010; Tremblay, Kraus, McGee, Ponton, & Otis 2001) have shown that the magnitude of this component increases with stimulus repetitions, suggesting that this component is sensitive to the statistics of the experiment. We hypothesized that P2's sensitivity to stimulus repetitions is a special case of its sensitivity to the (frequency) distance
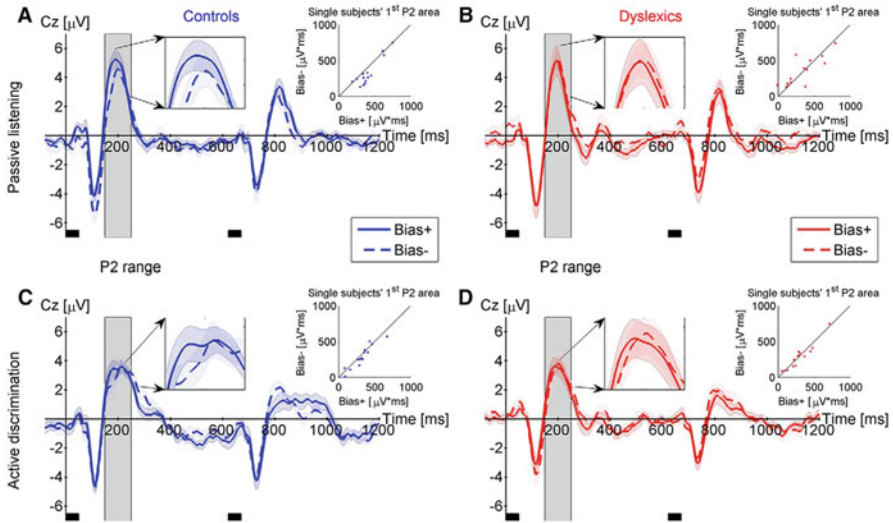
Fig. 10.6 Grand average ERP measures for the *Bias+* and *Bias−* trial types (electrode Cz). (**a, c**) controls (blue lines). (**b, d**) dyslexics (red lines). *Bias+* trials are denoted by solid lines and *Bias−* trials by dashed lines. In controls, the area of P2 after the first tone (from 150 to 250 ms, denoted by the gray rectangles) was significantly different between *Bias+* and *Bias−* trial types, in both passive listening (**a**) and during active discrimination (**c**) Dyslexics' evoked responses did not differ between the two trial types (**b, d**). Filled areas around the mean response denote cross-subject SEM. Small black rectangles under the plots denote the temporal location of the two tones in the trial. Insets: middle of each plot – P2 region enlarged; top right of each plot – single subject data of *Bias−* versus *Bias+* trials. In the Passive condition the difference between the trial types was significantly larger among controls than among dyslexics. (Adapted from Jaffe-Dax et al. 2015, with permission)

between the current stimulus and the mean of previous trials. Therefore, we predicted that the magnitude of control's P2 would be larger in *Bias+* trials than in *Bias−* trials, since the average distance of the first tone from the mean frequency is smaller in *Bias+* trials than in *Bias−* trials (as shown in Fig. 10.4). Consequently, the first tone in *Bias+* trials is closer to the mean (prior) than in *Bias−* trials.

We recorded ERPs with the same series of stimuli when participants either performed the task or watched a silent movie. For each participant in each of the specified trial types, we calculated the area under the curve between 150 and 250 ms after the first tone's onset as his/her individual P2 area. As predicted, we found that controls' automatically evoked response (Fig. 10.6a) was larger after the first tone in *Bias+* compared with *Bias−* trials. However, dyslexics' P2 was not sensitive to trial type, namely, *Bias+* and *Bias−* trials induced similar P2 components (Fig. 10.6b). Similar results were found under active (Fig. 10.6c,d) and passive (Fig. 10.6a, b) conditions. Taken together, these results support the hypothesis that dyslexics' computational deficit is associated with a failure to automatically integrate their on-line representations with the prior distribution.

10.7   Dyslexics' Underweight Prior also in the Visual Modality

Sensitivity to the statistics of the stimuli should not, theoretically, be restricted to a specific modality. We thus asked whether similarly reduced sensitivity would be found in the visual modality under statistically similar conditions. Previous work has shown that dyslexics' difficulty in visual discrimination tasks is restricted to sequential protocols, but their performance is intact on simultaneously presented stimuli (Ben-Yehudah & Ahissar 2004; Ben-Yehudah, Sackett, Malchi-Ginzberg, & Ahissar 2001). However, these visual studies only used typical protocols, in which one of the stimuli in each pair is constant across trials (reference) and the other (target) is randomly drawn from a limited range. Hence, serial comparisons could be aided by priors (references, anchors) based on previous trials, and consequently dyslexics' difficulties could be attributed to either poor explicit within-trial retention (explicit working memory), or to inefficient integration of priors (or both). Here (Jaffe-Dax, Lieder, Biron, & Ahissar 2016) we used the sequential spatial frequency discrimination task with a richer protocol where both gratings were randomly drawn from a wide range of spatial frequency (as in the auditory No-Reference protocol; Fig. 10.7). Thus, in addition to the groups' average performance, this protocol served to assess the magnitude of the contraction bias, which reflects the efficiency of using priors based on previous trials.

The two groups did not differ significantly in their Just Noticeable Differences (JNDs). To compare contraction bias, we compared individual performance in *Bias+* to performance on trials in *Bias−* range. Both populations showed contraction bias, i.e., better performance on trials when the first grating was closer to the mean frequency than the second grating, as shown in Fig. 10.8a. However, among dyslexics, the difference in performance between *Bias+* and *Bias−* trials was significantly smaller than among controls (Fig. 10.8a). This difference was consistent across participants: 82.5% of the control participants compared to only 57.5% of the dyslexic participants performed better on the Bias+ trials than on the Bias− trials (Fig. 10.8b).

10.8   Summary and Limitations

The deficits that we found characterize more than half of the dyslexics that we assessed, though not all of our participants. These deficits are expected to impede their ability to form robust representations of their native language statistics. Yet, the actual developmental trajectory of predictive abilities and linguistic development remains a topic for future research. Additionally, we have not addressed here the relations of the anchoring, or predictive coding hypothesis to other accounts of dyslexia. One hypothesis consistent with our results, is that dyslexics' deficits stem from some type of abnormality in the dorsal stream, whose impairment had
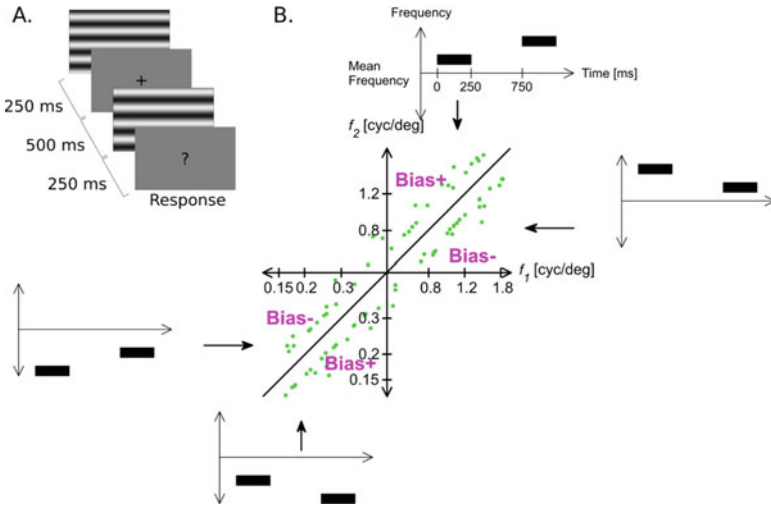
Fig. 10.7 Schematic illustrations of the sequential spatial frequency discrimination task and contraction towards the mean. (**a**) The temporal structure of a single trial. The first grating was presented for 250 ms, followed by an ISI of 500 ms. The second grating was presented for 250 ms. The observer was requested to indicate which of the two gratings "was denser" (had the higher spatial frequency). (**b**). The middle plot illustrates the distribution of single trials in the frequency plane (the frequencies of the first and second grating in each trial, respectively) for a typical subject. Each green dot denotes a pair of stimuli composing a single trial. This plane illustrates the ranges of the different trial types. Just as in the auditory frequency discrimination, in *Bias+* trials the frequency of the first grating stimulus was closer to the mean frequency; thus, contraction of its representation towards the mean increased the perceived difference between the two gratings and consequently improved performance. In *Bias−* trials the first grating was farther from the mean; thus, contraction of its representation towards the mean frequency decreased the perceived difference between the gratings and hampered performance. (Adapted from Jaffe-Dax et al. 2016, with permission)
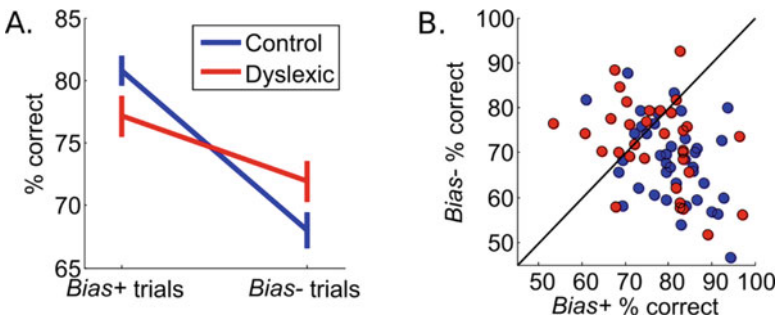

Fig. 10.8 Dyslexics' contraction bias in the visual domain is smaller than controls'. (**a**) Contraction bias averaged across participants. Ordinate shows the percentage of correct responses for the two sub-divisions of trials (abscissa): *Bias+* trials (left), and Bias− trials (right). Controls are denoted in blue, and dyslexics in red. Dyslexics' difference is smaller than controls', in spite of an overall similar % correct. Error bars denote SEM. (**b**) Individuals' performance (% accuracy) in *Bias−* versus *Bias+* trials. The diagonal indicates equal performance (no bias). Control participants (blue symbols) are distributed mainly below the diagonal, whereas dyslexic participants (red symbols) are more evenly distributed around the diagonal. (Adapted from Jaffe-Dax et al. 2016, with permission)

been previously suggested for dyslexia (Boros et al. 2016; Gori, Seitz, Ronconi, Franceschini, & Facoetti 2015; Paulesu, Danelli, & Berlingeri 2014). According to this interpretation, the dorsal stream is involved in serial processing in perception as well as in motor plans, and based on that, perhaps in serial planning and in working memory. In fact, recent imaging studies that have tried to dissociate the role of the dorsal and ventral streams in the context of speech perception suggest that the dorsal stream involves the fronto-parietal articulatory network, which is also related to working memory (Hickok & Poeppel 2015). The left fronto-parietal network is activated both in auditory 2-tone frequency discrimination tasks (Daikhin & Ahissar 2015), and in serial spatial frequency discrimination (Reinvang, Magnussen, & Greenlee 2002, though here in both the right and left hemispheres). One of the main bundles connecting posterior and frontal parts of the dorsal stream is the arcuate fasciculus (Dick & Tremblay 2012), whose abnormality in dyslexia has been suggested by previous studies (e.g., Boets et al. 2013; Klingberg et al. 2000). These might be the candidates for the neurological origin of the deficient statistical learning in dyslexia.

References

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387*(6631), 401–406. https://doi.org/10.1038/387401a0

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8*(10), 457–464. https://doi.org/10.1016/j.tics.2004.08.011

Ahissar, M., Lubin, Y., Putter-Katz, H., & Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nature Neuroscience, 9*(12), 1558–1564. https://doi.org/10.1038/nn1800

Ahissar, M., Protopapas, A., Reid, M., & Merzenich, M. M. (2000). Auditory processing parallels reading abilities in adults. *Proceedings of the National Academy of Sciences of the United States of America, 97*(12), 6832–6837. https://doi.org/10.1073/pnas.97.12.6832

Amitay, S., Ahissar, M., & Nelken, I. (2002). Auditory processing deficits in reading disabled adults. *Journal of the Association for Research in Otolaryngology, 3*(3), 302–320. https://doi.org/10.1007/s101620010093

Ashourian, P., & Loewenstein, Y. (2011). Bayesian inference underlies the contraction bias in delayed comparison tasks. *PloS One, 6*(5), e19551. https://doi.org/10.1371/journal.pone.0019551

Baldeweg, T. (2007). ERP repetition effects and mismatch negativity generation. *Journal of Psychophysiology, 21*(3–4), 204–213. https://doi.org/10.1027/0269-8803.21.34.204

Baldo, J. V., & Dronkers, N. F. (2006). The role of inferior parietal and inferior frontal cortex in working memory. *Neuropsychology, 20*(5), 529–538. https://doi.org/10.1037/0894-4105.20.5.529

Banai, K., & Ahissar, M. (2017). Poor sensitivity to sound statistics impairs the acquisition of speech categories in dyslexia. *Language, Cognition and Neuroscience, 11*(11), 1–12. https://doi.org/10.1080/23273798.2017.1408851

Ben-Yehudah, G., & Ahissar, M. (2004). Sequential spatial frequency discrimination is consistently impaired among adult dyslexics. *Vision Research, 44*(10), 1047–1063. https://doi.org/10.1016/j.visres.2003.12.001

Ben-Yehudah, G., Sackett, E., Malchi-Ginzberg, L., & Ahissar, M. (2001). Impaired temporal contrast sensitivity in dyslexics is specific to retain-and-compare paradigms. *Brain, 124*(7), 1381–1395. https://doi.org/10.1093/brain/124.7.1381

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex, 10*(5), 512–528.

Boets, B., Op de Beeck, Hans P, Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., …Ghesquière, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science, 342*(6163), 1251–1254. https://doi.org/10.1126/science.1244333

Boros, M., Anton, J.-L., Pech-Georgel, C., Grainger, J., Szwed, M., & Ziegler, J. C. (2016). Orthographic processing deficits in developmental dyslexia: Beyond the ventral visual stream. *NeuroImage, 128*, 316–327. https://doi.org/10.1016/j.neuroimage.2016.01.014

Bull, A. R., & Cuddy, L. L. (1972). Recognition memory for pitch of fixed and roving stimulus tones. *Perception & Psychophysics, 11*(1), 105–109. https://doi.org/10.3758/BF03212696

Cohen, Y., Daikhin, L., & Ahissar, M. (2013). Perceptual learning is specific to the trained structure of information. *Journal of Cognitive Neuroscience, 25*(12), 2047–2060. https://doi.org/10.1162/jocn_a_00453

Daikhin, L., & Ahissar, M. (2015). Fast learning of simple perceptual discriminations reduces brain activation in working memory and in high-level auditory regions. *Journal of Cognitive Neuroscience, 27*(7), 1308–1321. https://doi.org/10.1162/jocn_a_00786

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *The Journal of Neuroscience, 23*(8), 3423–3431.

Dick, A. S., & Tremblay, P. (2012). Beyond the arcuate fasciculus: Consensus and controversy in the connectional anatomy of language. *Brain, 135*(12), 3529–3550. https://doi.org/10.1093/brain/aws222

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. https://doi.org/10.1016/j.tics.2010.01.004

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*(10), 475–483. https://doi.org/10.1016/S0166-2236(00)01633-7

Friederici, A. D., Makuuchi, M., & Bahlmann, J. (2009). The role of the posterior superior temporal cortex in sentence comprehension. *Neuroreport, 20*(6), 563–568. https://doi.org/10.1097/WNR.0b013e3283297dee

Gori, S., Seitz, A. R., Ronconi, L., Franceschini, S., & Facoetti, A. (2015). Multiple causal links between magnocellular-dorsal pathway deficit and developmental dyslexia. *Cerebral Cortex, 26*(11), 4356–4369. https://doi.org/10.1093/cercor/bhv206

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *The Journal of Neuroscience, 25*(45), 10494–10501. https://doi.org/10.1523/JNEUROSCI.1227-05.2005

Hickok, G., & Poeppel, D. (2015). Chapter 8: Neural basis of speech perception. In M. J. Aminoff, F. Boller, & D. F. Swaab (Eds.), *The human auditory system* (pp. 149–160). Amsterdam: Elsevier.

Hollingworth, H. L. (1910). The central tendency of judgment. *The Journal of Philosophy, Psychology and Scientific Methods, 7*(17), 461. https://doi.org/10.2307/2012819

Jaffe-Dax, S., Frenkel, O., & Ahissar, M. (2017). Dyslexics' faster decay of implicit memory for sounds and words is manifested in their shorter neural adaptation. *eLife, 6*, e20557. https://doi.org/10.7554/eLife.20557

Jaffe-Dax, S., Lieder, I., Biron, T., & Ahissar, M. (2016). Dyslexics' usage of visual priors is impaired. *Journal of Vision, 16*(9), 10. https://doi.org/10.1167/16.9.10

Jaffe-Dax, S., Raviv, O., Jacoby, N., Loewenstein, Y., & Ahissar, M. (2015). A computational model of implicit memory captures dyslexics' perceptual deficits. *The Journal of Neuroscience, 35*(35), 12116–12126. https://doi.org/10.1523/JNEUROSCI.1302-15.2015

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar Straus and Giroux.

Klingberg, T., Hedehus, M., Temple, E., Salz, T., Gabrieli, J. D., Moseley, M. E., & Poldrack, R. A. (2000). Microstructure of temporo-parietal white matter as a basis for reading ability: Evidence from DTI. *European Psychiatry, 17*, 48. https://doi.org/10.1016/S0924-9338(02)80215-2

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Muller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping, 30*(3), 859–873. https://doi.org/10.1002/hbm.20550

Magen, H., Emmanouil, T.-A., McMains, S. A., Kastner, S., & Treisman, A. (2009). Attentional demands predict short-term memory load response in posterior parietal cortex. *Neuropsychologia, 47*(8–9), 1790–1798. https://doi.org/10.1016/j.neuropsychologia.2009.02.015

Mayhew, S. D., Dirckx, S. G., Niazy, R. K., Iannetti, G. D., & Wise, R. G. (2010). EEG signatures of auditory activity correlate with simultaneously recorded fMRI responses in humans. *NeuroImage, 49*(1), 849–864. https://doi.org/10.1016/j.neuroimage.2009.06.080

Nahum, M., Daikhin, L., Lubin, Y., Cohen, Y., & Ahissar, M. (2010). From comparison to classification: A cortical tool for boosting perception. *The Journal of Neuroscience, 30*(3), 1128–1136. https://doi.org/10.1523/JNEUROSCI.1781-09.2010

Obleser, J., & Kotz, S. A. (2010). Expectancy constraints in degraded speech modulate the language comprehension network. *Cerebral Cortex, 20*(3), 633–640. https://doi.org/10.1093/cercor/bhp128

Oganian, Y., & Ahissar, M. (2012). Poor anchoring limits dyslexics' perceptual, memory, and reading skills. *Neuropsychologia, 50*(8), 1895–1905. https://doi.org/10.1016/j.neuropsychologia.2012.04.014

Paulesu, E., Danelli, L., & Berlingeri, M. (2014). Reading the dyslexic brain: Multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Frontiers in Human Neuroscience, 8*, 830. https://doi.org/10.3389/fnhum.2014.00830

Preuschhof, C., Schubert, T., Villringer, A., & Heekeren, H. R. (2010). Prior information biases stimulus representations during vibrotactile decision making. *Journal of Cognitive Neuroscience, 22*(5), 875–887. https://doi.org/10.1162/jocn.2009.21260

Raviv, O., Ahissar, M., & Loewenstein, Y. (2012). How recent history affects perception: The normative approach and its heuristic approximation. *PLoS Computational Biology, 8*(10), e1002731. https://doi.org/10.1371/journal.pcbi.1002731

Reinvang, I., Magnussen, S., & Greenlee, M. W. (2002). Hemispheric asymmetry in visual discrimination and memory: ERP evidence for the spatial frequency hypothesis. *Experimental Brain Research, 144*(4), 483–495. https://doi.org/10.1007/s00221-002-1076-y

Rokem, A., & Ahissar, M. (2009). Interactions of cognitive and auditory abilities in congenitally blind individuals. *Neuropsychologia, 47*(3), 843–848. https://doi.org/10.1016/j.neuropsychologia.2008.12.017

Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research, 248*(1–2), 48–59. https://doi.org/10.1016/j.heares.2008.11.012

Sheehan, K. A., McArthur, G. M., & Bishop, D. V. M. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Brain Research. Cognitive Brain Research, 25*(2), 547–553. https://doi.org/10.1016/j.cogbrainres.2005.08.007

Sreenivasan, K. K., Curtis, C. E., & D'Esposito, M. (2014). Revisiting the role of persistent neural activity during working memory. *Trends in Cognitive Sciences, 18*(2), 82–89. https://doi.org/10.1016/j.tics.2013.12.001

Tong, Y., Melara, R. D., & Rao, A. (2009). P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Research, 1297*, 80–88. https://doi.org/10.1016/j.brainres.2009.07.089

Tremblay, K., Inoue, K., McClannahan, K., & Ross, B. (2010). Repeated stimulus exposure alters the way sound is encoded in the human brain. *PloS One, 5*(4), e10283. https://doi.org/10.1371/journal.pone.0010283

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. *Ear and Hearing, 22*(2), 79–90. https://doi.org/10.1097/00003446-200104000-00001

Wickelgren, W. A. (1969). Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology, 6* (1), 13–61. https://doi.org/10.1016/0022-2496(69)90028-5

Woodrow, H. (1933). Weight-discrimination with a varying standard. *The American Journal of Psychology, 45*(3), 391. https://doi.org/10.2307/1415039

Woolgar, A., Parr, A., Cusack, R., Thompson, R., Nimmo-Smith, I., Torralva, T., . . . Duncan, J. (2010). Fluid intelligence loss linked to restricted regions of damage within frontal and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 107*(33), 14899–14902. https://doi.org/10.1073/pnas.1007928107

World Health Organization (Ed.). (2008). *International statistical classification of diseases and related health problems* (10 ed.). Berlin/Heidelberg: Springer. https://doi.org/10.1007/SpringerReference