Beyond DSM: The Role of Auditory Processing in Attention and Its Disorders

Teresa Bailey

Private Practice, Los Altos, California and Department of Research, Athena Academy, Palo Alto, California

To cite this article: Teresa Bailey (2012): Beyond DSM: The Role of Auditory Processing in Attention and Its Disorders, Applied Neuropsychology: Child, 1:2, 112-120

To link to this article: http://dx.doi.org/10.1080/21622965.2012.703890
Beyond DSM: The Role of Auditory Processing in Attention and Its Disorders

Teresa Bailey

Private Practice, Los Altos, California and Department of Research, Athena Academy, Palo Alto, California

This article reviews and synthesizes recent research regarding auditory processing, attention, and their roles in generating both adaptive and maladaptive behavioral responses. Research in these areas is beginning to converge on the role of polymorphisms associated with catecholamine metabolism and transport, particularly the neurotransmitter dopamine. The synthesis offered in this article appears to be the first to argue that genetic differences in dopamine metabolism may be the common factor in four disparate disorders that are often observed to be comorbid, i.e., attention-deficit hyperactivity disorder, auditory processing disorders, developmental language disorders, and reading disorders.

Key words: attention, attention-deficit hyperactivity disorder, auditory processing, dopamine, dyslexia, ERP, event-related potential, genetics, language, mismatch negativity, reading

According to the Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association, 2000), a diagnosis of attention-deficit hyperactivity disorder (ADHD)-Combined type can be made from a combination of 18 unprioritized symptoms evenly divided between two clusters of 9 each. Of the 18 symptoms, 2—both in the Inattention cluster—make reference to the role of sensory input and behavioral response. The first is a negative behavior (i.e., a failure to respond): 1(c) “often does not seem to listen when spoken to directly.” The second, 1(h), is an overresponse to the environment: “is often easily distracted by extraneous stimuli” (p. 65). There is no requirement for a diagnostic investigation or verification of the underlying cause(s) of failure(s) to listen or the reason(s) for the easily distracted behaviors, except that they not be due to a pervasive developmental, psychotic, anxiety, or mood disorder.

Regardless of the clinical dissatisfactions with the DSM behavioral checklist approach to diagnosis in an age of neuroscience’s ability to measure many underlying components of the final common pathways of observable behaviors, two facts remain: Many individuals respond well to stimulant medications, at least for a time, and the relative success of that treatment model compared with other mainstream interventions is based on a theory of attention, concentration, and impulse regulation strongly influenced by a dopamine theory of attention (Molina et al., 2009).

This article will focus on poor auditory processing as a possible factor that contributes to the diagnosis and treatment of inattentive behaviors. In addition, the mutual influences between auditory processing and attention, and the role they play in generating motor activity will place the diagnosis and misdiagnosis of ADHD as a behaviorally defined syndrome in a neurodevelopmental neurobehavioral context.

Part of the confusion about the role of auditory processing in attention disorders has to do with which professional a family first turns to in assistance with the diagnosis. If the presenting problem is poor listening,
Types of Attention

Coherent discussion of attention deficits is difficult because there is no unitary definition of attention. There are subtypes that are not empirically defined, and the boundaries between them can be imprecise. When does arousal become orientation, and orientation become focus? These are neurophysiologic questions that will likely be answered more precisely with further investigation. When is focus lost or maintained while “shifting set,” or manifested as divided attention? When does rapid set shifting become distractibility? Outside of scanning an environment for basic nourishment and safety, these distinctions are largely culturally determined judgments of relative performance levels in a specific setting, within a range of tolerances that are often dependent on the setting itself. For example, different scanning and social interaction behaviors are expected and required in a classroom lecture than are those expected and required in a large, noisy wedding reception. What counts as distracted behavior in one setting may or may not be acceptable, or even expected, in another. A discussion of attention subtypes is beyond the scope of this article. The focus here will be on the close interaction among auditory processing, various aspects of attention, and behavioral responses to processed auditory sensory stimuli.

COMPLEX INTERACTIONS AMONG AUDITORY PROCESSING, ATTENTION, AND BEHAVIORAL RESPONSES

Auditory Processing is More Than Hearing

Poor listening is a common symptom in relatively narrow evaluations of ruling in or ruling out a disorder of attention, as well as within a comprehensive evaluation of multiple neurodevelopmental and behavioral problems. While there is widespread understanding that language facility and sophistication develops over many years, there is less general awareness that auditory function in children is not the same as that of adults. When a child passes screening measures for peripheral hearing, it is largely assumed that any failure to listen or to understand must be either ADHD or poor motivation to listen. Hearing is established before birth, and parts of the auditory system are relatively mature at birth (Hall, 2007; Picton & Taylor, 2007). Other auditory functions such as accurate sound discrimination in reverberant noisy rooms do not reach mature levels until the teen years (Werner, 2007). Decades of research have demonstrated that various aspects of central auditory processing have a measurable developmental trajectory (Berlin, Hughes, Lowe-Bell, & Berlin, 1973; Birkas et al., 2006; D’Angiulli, Herdman, Stapells, & Hertzman, 2008; de Bie et al., 2012; Gomes, Duff,
both ipsilateral and bilateral subcortical processing of auditory signal processing as part of distinct and interdependent networks is crucial for the eventual cognitive (Dawes & Bishop, 2009; Rosen, Cohen, & Vanniasegaram, 2010) and behavioral response of the individual to environmental demands. Extending these findings into the role of auditory processing in the development of language skills relating to interpersonal social relations, it becomes clear that part of the child’s ability to participate accurately and efficiently in social language can greatly affect that individual’s sense of efficacy in the social environment.

Just as dopamine plays a role in the regulation of attention and impulsivity, significant components of auditory processing are mediated by dopamine in the basal ganglia–frontal connections and in cerebellar reciprocal loops (Birkas et al., 2006; Kashino & Kondo, 2012; Majic et al., 2010; Sens, de Almeida, de Souza, Goncalves, & do Carmo, 2011), as well as within the auditory cortex itself (Birkas et al.). Functional inefficiencies in auditory processing result in behaviors that are the clinical focus of neuropsychological assessment due to failures of a child to adapt to common situational demands across settings. Among these adaptive failures are difficulty with listening to and understanding oral instructions and an inability or variable ability to respond adequately within expected time frames, with a level of focus and behavior appropriate to the situation.

From Hearing to Listening to Behavioral Response

Auditory attention is highly dependent on intact peripheral hearing. If the incoming sensory signal is degraded due to problems in the peripheral system, it follows that there will be suboptimal processing and behavioral response. The American Speech Language Hearing Association (2005) and the American Academy of Audiology (2010) both require that there be a demonstration of intact peripheral hearing before a primary auditory processing disorder can be diagnosed. Most children presenting for neuropsychological evaluation have passed a hearing screening at school and/or their pediatrician’s office. Information about a history of otitis media should be obtained. Most cases resolve adequately, but the prolonged loss of clear signals in the early years of language development can result in disruptions of language and listening skills that may not be absolute, but that can result in subtle listening problems (Maruthy & Mannurukrishnaiah, 2008; Mody et al., 1999; Zumach et al., 2009).

Once an auditory signal has been received, there is both ipsilateral and bilateral subcortical processing of components of the signal, including pitch or frequency as measured in Hertz, intensity or loudness, and duration. These characteristics of speech and nonspeech sounds are further sorted and salience is weighted throughout the brainstem, other subcortical structures, including the cerebellum, on up to the cortical level. Afferent and efferent single pathways and reciprocal feedback loops combine into several parallel and integrated networks. Based on millisecond interaural timing differences beginning at the superior olivary complex, a determination of the sound source location is calculated through complex interactions of subcortical and cerebellar bidirectional feedback loops (Bailey, 2010; Musiek & Baran, 2007; Musiek & Chermak, 2007). This information is combined with interpretive decisions about the meaning of the sound, based on processes that are still not well understood. The efficiency of this decision process is, in part, based on an individual’s developmental level and cumulative experience (Herdman, 2011), as well as genetic variations (Kashino & Kondo, 2012; Majic et al., 2011). Once a saliency determination is made, a behavioral response will be generated. The behavioral options are to do something different (e.g., move toward or away from the sound source [shifting attention with a change in behavior]), decide the sound is irrelevant and continue with what the individual was already doing [focused or sustained attention], or a compromise of continuing with the current activity but continuing to monitor the environment [divided attention]). In this way, sensory processing in general and auditory processing in particular are involved with environmental scanning for the purposes of behavioral response.

Behavioral responses often include integration and coordination with other sensory inputs, such as the visual system, to guide the resulting motor behaviors. Motor behavior as a response to auditory signaling must then be coordinated and continually updated through complex afferent and efferent networks (Arnott & Alain, 2011; Nääätänen, Kujala, & Winkler, 2011; Salmi, Rinne, Degerman, Salonen, & Alho, 2007; Shiels & Hawk, 2010).

The ability to form age-appropriate behavioral responses to environmental demands depends on the ability to interpret accurately the demands themselves. These demands often have a significant auditory component, either speech or other environmental sounds such as sirens, automobiles, or animals. The role of the clinician is to distinguish between a child’s ability to listen and possible refusal to comply, and a child’s inability to interpret accurately and understand what is said, as the source of the problem that prevents compliance. If peripheral hearing is normal and there is no significant delay in language, the options for understanding the observed behaviors in most children, until recently, have
been ADHD or emotional factors. An increase in clinical awareness of the developmental trajectories in auditory processing abilities may add to clinical understanding by explaining an otherwise obscured portion of the underlying reason(s) about why a child may be inattentive and develops negative emotional responses to an inability to attend, understand, and respond adequately to a situation.

BEYOND BEHAVIORAL CHECKLISTS: EMPIRICAL EVIDENCE FOR AUDITORY PROCESSING, ATTENTION, AND THEIR CONFLUENCE

Research in evoked-response potentials (ERPs) has significantly furthered understanding of the intersection between sensory processing and the role of attentional shifts in detecting novel information within an otherwise habituated sensory stream (Hall, 2007; Sussman & Steinschneider, 2009; Thompson & Spencer, 1966). Electroencephalogram (EEG) research has seen a revival because it has millisecond temporal resolution, compared with limited functional magnetic resonance imaging (fMRI) temporal resolution of approximately 1 second. This is at least two orders of magnitude slower than recordings made with EEG sampling. Furthermore, unlike EEG, fMRI is unable to account for variability from trial to trial or to indicate the directionality of information transmission within a network (Smith, Pillai, Chen, & Horwitz, 2012). EEG auditory ERPs reveal a developmentally normative set of peaks and valleys within distinct millisecond time windows as the signal passes through subcortical to cortical structures. In addition, deviations from these normative waveforms, both their amplitudes and time windows, can give clinically useful information about processing efficiency and attending behaviors in both auditory and visual modalities (Hall).

Gomes et al. (2007) demonstrated the wave morphological development of the Nd and P3b waves for selective auditory attention in normally developing 9- and 12-year-old children. They compared the children’s waveforms with mature adult waveforms. Figure 1 shows the means of the ERP waveforms for both attended and nonattended nonspeech tones along the anterior–posterior central axis. The real-time measurement of EEG functions demonstrates the difference between receptive areas (posterior; Pz), interpretive, central areas (central; Cz), and the eventual arrival at frontal areas for behavioral decision-making (frontal;
Fz) and motor planning and execution (FCz). Here the simplification and flattening of the mature waves may indicate an objective representation of the automation of experiential learning. The simplification of the waves may represent increased efficiency due to more automatic behavioral choices having been moved from cortical to subcortical response patterns (Herdman, 2011; Näätänen et al., 2011). Further interpretation of these forms is beyond the scope of this article. A thorough introduction can be found in Hall (2007), and a brief introduction can be found in Bailey (2010). In general, these waves indicate that there is an objectively measurable, normative developmental trajectory for aspects of auditory processing (Räikkönen, Birkas, Horvath, Gervai, & Winkler, 2003).

How this information is conveyed through complex networks and the decisions made about whether, and if so, which and how much motor response to make, have been the subject of intense investigation for more than 30 years. A core area of research has been to understand attended and unattended ERPs. After these data are obtained, a difference wave is calculated between the two signal conditions, called mismatch negativity (MMN). Näätänen et al. (2011) demonstrated how MMN allowed for the investigation of the subcomponents of the N1 (negative wave at approximately 100 milliseconds poststimulus onset) and PN (the resulting positive wave after the N1 wave) indices of selective attention. These findings were correlated and confirmed with magnetoencephalographic and fMRI studies.

Näätänen et al. (2011) reviewed and extended their findings to create a more sophisticated explanation of sensory monitoring, sensory processing, attention, and motor response to changes in the environment compared with those previously described primarily as bottom-up versus top-down models. They accomplished this through analysis of the subcomponents of waves generated by attended and unattended auditory signals that form the MMN wave. They proposed that the MMN wave and its subcomponents represent an indication of the automatic switching of attention when there is a change in the overall auditory scene. Switching is based on the brain’s ability to identify and orient to new auditory stimuli. The N1 component wave detects the presence of a sound. The MMN detects the difference between the ambient auditory scene and the newly detected sound(s). The PN wave morphology is affected by selected attention to one sound versus another. After detection of a deviant, unexpected sound, the sounds are processed through various unidirectional and bidirectional networks across cortical, subcortical, and cerebellar...

![Figure 2](attachment:image_url)  
**Figure 2.** A model of conscious and unconscious processes in audition. The sound stimulus is first very rapidly analyzed by the different feature detectors. Thereafter, the outputs from the different feature detectors are temporarily integrated with each other in the temporal window of integration. The accumulation of this integrated sensory information in the mechanisms of sensory memory that evolve in time provides the sensory data of subjective contents of percepts (i.e., the central sound representation; Näätänen & Winkler, 1999). This central representation becomes consciously experienced, depending on the strength of the attention-call signal elicited by the dynamogenic stimulus features indexed by the N1 amplitude. Further, if some discernible change in auditory stimulation occurs, then this change results in the updating of auditory representations in sensory memory, eliciting the auditory-cortex MMN component. This, in turn, activates the frontal cortex mechanisms generating the frontal MMN component (representing an attention-call signal to auditory change). During selective attention, the executive mechanisms use fresh sensory memory data to set up and tune the attentional trace, a temporary template for the rapid selection of the to-be-attended input for further processing or response. This selection mechanism continuously depends on the active maintenance and rehearsal of the aspects of sensory input that very rapidly enable the listener to distinguish the relevant sensory input stream among the concurrent stimulus streams.
circuit, according to various characteristics. These include but are not limited to pitch, duration, intensity, whether or not it is speech, and saliency (Kashino & Kondo, 2012). Because one of the frontal functions is to control where attention is directed in order to generate a behavioral response, good auditory processing is important for efficient frontal functioning. Lesion studies show that dorsolateral prefrontal lesions attenuate MMN amplitudes (Alain, Woods, & Knight, 1998; Alhol, Woods, Algazi, Knight, & Näätänen, 1994). Näätänen et al. proposed a dynamic model based on the accumulated literature of these investigations. Figure 2 shows functional integration of auditory sensory processing, attention, and working memory for the purposes of generating a behavioral response.

Despite the progress that has been made, some controversy remains regarding how best to measure auditory ERPs and how to best make clinical diagnoses based on these findings. Many ERP studies involve presentation of nonspeech tones or clicks. There is another literature using nonsense consonant–vowel or consonant–vowel–consonant single-syllable words. Tones and clicks have the advantage of being relatively free of language, accent, and cultural factors. In light of these difficulties, some researchers are beginning to argue that auditory processing should only be assessed with nonspeech stimuli (Dawes & Bishop, 2009). Kashino and Kondo (2012) emphasize, however, that actual perception–interpretation–action sequences involve monitoring and response to an ever-shifting auditory scene. The scene includes a mix of speech and nonspeech auditory signals that require, and elicit, different neural responses depending on the type of sound. There are specific receptors throughout the auditory and language areas for speech that do not respond to nonspeech sounds.

THE GENETICS OF DOPAMINE-RELATED
POLYMORPHISMS HELP EXPLAIN WHY
PROBLEMS WITH ATTENTION, IMPULSIVITY,
AUDITORY PROCESSING, AND READING
ARE OFTEN COMORBID

In a complex study of the genetics of the catecholamines dopamine and noradrenalin and fMRI evidence for perceptual networks, Kashino and Kondo (2012) began to construct a neural basis of the intersection among attention, auditory processing, and verbal language processing. They used fMRI to construct several neural networks. These networks were related to speech and nonspeech sound processing, orientation to new sounds, shifting attention based on linguistic and nonlinguistic features, and the efficiency and stability of attention shifts based on polymorphisms of the COMT Val 158Met gene. The COMT Val 158Met gene is significantly involved in the degradation of catecholamines such as dopamine and noradrenalin. Dopamine plays a central role in the regulation of impulsive motor urges and actions as well as attention. Using auditory scene interpretation, including both accuracy and efficiency, Kashino and Kondo demonstrated that genetic differences in COMT genes might account for individual variations of the balance between exploring novel perceptions in the environment and the need for stability and focus. They proposed a model of neural networks that involve predicting and interpreting novel auditory stimuli that involves motor prediction and calibration in the cerebellum, the basal ganglia, frontal regions, and the auditory cortex as part of the phonetic and rhythmic aspects of acoustic sequences. They suggested that this balance is modulated by both dopamine and noradrenalin.

Dopamine has been found in afferent and efferent auditory pathways that presumably connect with structures that are involved in motor decisions (Kudoh & Shibuki, 2006; Ruel et al., 2001, 2006). The complex interaction among auditory sensory input from the environment, attention, and impulse management has only recently begun to be investigated as a network of mutual neural influences (Andersson, Llera, Rimol, & Hugdahl, 2008; Kashino & Kondo, 2012; Näätänen et al., 2011; Shiels & Hawk, 2010).

Willcutt, Pennington, Olson, and DeFries (2007) have synthesized a large body of work on heritability and comorbidity of ADHD and reading disorders. They have found that polymorphisms in dopamine genes account for a wide variability in ADHD and reading disability presentations. Their studies suggest that the genetic loading accounts for approximately 70% of the variance seen in ADHD, although no one dopamine-related gene appears to exert a controlling influence compared with other genes that encode for dopamine metabolism and transport. Dopamine is being shown to have a role in auditory processing (Kudoh & Shibuki, 2006; Ruel et al., 2006). It appears that a possible resolution to the confusion and controversy about the overlap among ADHD, language-based learning disabilities (Bishop, Adams, & Norbury, 2006) including dyslexia (Willcutt et al.) and auditory processing disorders may be found in further investigations of dopamine regulation that variously affects all four disorders.

These often comorbid disorders, however, are developmental in nature. Their complex genetic components interact with individual histories and environments. Some aspects of language development, including many aspects of grammar, are considered to be mature by 4 years of age. Reading, as a particular cultural task of language development, tends to be introduced at around 5 or 6 years of age. Bishop et al. (2006) demonstrated a genetic basis for phonological short-term
memory as shown by abilities for nonword repetition based on studies of 6-year-olds. They further argued that some differences in genes may be markers of language impairment. Sentence repetition did not show any genetic influence, even though it was able to differentiate high- and low-risk children regarding language impairment. Although their twin studies showed significant genetic effects, Bishop et al. (2006) did not identify any specific genes or polymorphisms in their study. Bishop et al. (1999) presented evidence that auditory processing ability is not heritable. More recent evidence from a twin study indicates that dichotic listening ability is highly heritable and accounts for approximately 70% of the variance in performance (Morrell et al., 2007). When they removed 12 twin pairs from the analysis, among whom 18 individuals met the diagnostic criteria for ADHD, their results did not change. Studies such as that by Morrell et al. demonstrate the importance of the choice of which aspect of auditory processing is measured and how it is measured. These choices greatly affect the kind of data that are generated in this field. Morell et al. found, as did Willcutt et al. (2007), that genes contribute approximately 70% of the variance and environment provides approximately 30%. Morrell et al. found this proportion to be comparable to other heritable conditions and traits such as insulin-dependent diabetes ($h^2 = .70$; Kyvik, Green, & Beck-Nielsen, 1995), height ($h^2 = .68$; Estourgie-van Burk, Bartels, van Beijsterveldt, Delemarre-van de Waal, & Boomsma, 2006), and phonological processing ($h^2 = .71$; Bishop et al., 1999).

**CLINICAL IMPLICATIONS OF AUDITORY PROCESSING FOR ASSESSING DISORDERS OF ATTENTION**

Because the diagnosis of ADHD requires symptoms that represent consistently inappropriate behavioral responses to environmental demands, it is crucial for clinicians to understand more about how and where the breakdowns contributing to the disordered behaviors have likely occurred so that treatment recommendations can be made to address the point(s) of breakdown more efficiently and effectively. To the extent that stimulant medication results in improvement in some cases but does not result in optimized behavior at the level of unaffected persons (Molina et al., 2009), it is incumbent upon clinicians to understand what combination of factors could be causing the behaviors and to redirect or broaden our focus to incorporate those potential causes into our case formulations and recommendations.

A full review of issues surrounding evaluation of auditory processing, especially by behavioral measures alone, is beyond the scope of this article. Dawes and Bishop (2009) have written a review on advantages and disadvantages of the auditory screening battery SCAN (Keith 2009a, 2009b) and other screening measures. They identify important shortcomings, including the SCAN battery only being valid for those with American-accented English, the sometimes complex instructions that increase the load for attention and language processing, and some retest reliability issues. Nevertheless, they argue in favor of the clinical usefulness of these screening measures as more useful than not for the added information they provide about refining diagnostic hypotheses about the sources of difficulties with attention, listening, language-based learning, and behavior regulation. In contrast, auditory ERP appears to be highly reliable and stable in healthy children as young as 6 years old retested at a 3-month interval (Rääkkönen et al., 2003).

Several studies address the use of dichotic listening tests in various formats as being particularly useful in the description of developmental trajectories and identification of auditory processing disorders (Berlin et al., 1973; Gomes et al., 2007; Sequeira, Specht, Moosmann, Westerhausen, & Hugdahl, 2010; Tillery et al., 2000; Westerhausen et al., 2011; Wexler & Halwes, 1985). Kudo and Shibuki (2006) emphasize the need to assess multiple aspects of listening, not just with single tones or words, because the individual who may be able to perform shorter tasks must also be able to interpret running speech in a variety of background auditory scenes.

Finally, it is increasingly being shown that there is a strong genetic loading for ADHD and reading disorders, both as individual and as comorbid disorders (Willcutt et al., 2007). Studies also have found a strong genetic influence for the development of grammar and language processing (Bishop et al., 2006). Some aspects of auditory processing appear to be strongly heritable (Morell et al., 2007). The heritability of aspects of auditory processing, like that of ADHD and reading disorders, appears to be reliably linked to polymorphisms of genes that encode for dopamine metabolism and transport (Kashino & Kondo, 2012; Kudo & Shibuki, 2006; Ruel et al., 2001, 2006). Research regarding the nature and extent of the mutual influences of these genes in constructing the neural networks that respond to stimuli, interpret and weigh the constantly shifting data streams, and produce both adaptive and maladaptive behaviors is finally being freed from top-down versus bottom-up dichotomous thinking and is entering a phase of documenting the complex interactions that more accurately reflect clinical observations and everyday experiences across multiple settings.
REFERENCES


Morrell, R. J., Brewer, C. C., Ge, D., Snieder, H., Zalewski, C. K., King, K. A., ... Friedman, T. B. (2007). A twin study of auditory processing indicates that dichotic listening ability is a strongly heritable trait. *Human Genetics*, 122, 103–111.


Niätäinen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: A unique window to central...


