Why does Music Therapy help in Autism?

NEHA KHETRAPAL[1]

University of Bielefeld, Germany

ABSTRACT: Music therapy is shown to be an effective intervention for emotional recognition deficits in autism. However, researchers to date have yet to propose a model that accounts for the neurobiological and cognitive components that are responsible for such improvements. The current paper outlines a model whereby the encoding of tonal pitch is proposed as the underlying mechanism. Accurate tonal pitch perception is important for recognizing emotions like happiness and sadness in the auditory domain. Once acquired, the ability to perceive tonal pitch functions as a domain-specific module that proves beneficial for music cognition. There is biological preparedness for the development of such a module and it is hypothesized to be preserved in autism. The current paper reinforces the need to build intervention programs based on this preserved module in autism, and proposes that this module may form the basis for a range of benefits related to music therapy. Possible brain areas associated with this module are suggested.

Submitted 2008 November 7; accepted 2009 January 12.

KEYWORDS: emotional processing, rehabilitation, music cognition, neurobiology

AUTISM

Kanner in 1943 was the first to describe and label autistic disorder. Autistic disorder is a neurodevelopmental condition marked by social and communication problems as well as restricted interests and behaviors (American Psychiatric Association [APA], 1994). It falls under the rubric of autistic spectrum disorder (ASD). Autism is usually diagnosed before three years of age and has a lifelong persistence. Males are usually more vulnerable with a ratio of 4 to 1 (Fombonne, 2005a). The prevalence of the disorder is approximately 13 in every 10,000 births (Fombonne, 2005b). The diagnostic and statistical manual, 4th edition, (DSM-IV; American Psychiatric Association) treats this disorder as a part of the category of Pervasive Developmental Disorder, which is characterized by three important disabilities: failure to develop normal social interaction; language delay and communication disability; and restricted, repetitive and stereotyped behaviors. Roughly 10% of ASD individuals show good rote memory and possess isolated savant skills like exceptional musical or drawing abilities or calendar calculations (Volkmar & Pauls, 2003).

Music Therapy as Treatment

Individuals with autism show poor perception of affective cues within the social domain and experimental evidence has also shown that such individuals often fail to interpret and recognize vocal and facial expressions of emotions (Capps, Yirmiya & Sigman; 1992; Fein, Lucci, Braverman & Waterhouse, 1992). Because the accurate perception and interpretation of such affective cues is so important for adaptive living, rehabilitation efforts have focused on ways in which these deficits can be alleviated. Music therapy, as a result, has been employed with success in autism (Kaplan & Steele, 2005). This is likely because individuals with autism show no deficits in processing musical affect (Heaton, Hermelin & Pring, 1999).

There are two major goals when music therapy is employed for individuals with autism: improving communication/language and improving socio-behavioral skills (Kaplan & Steele, 2005). For instance, Edgerton (1994) investigated the effect of improvisational music therapy on communicative behaviors and found gains for musical communicative modalities (tempo, rhythm, form
and pitch) across ten music therapy sessions. Furthermore, a significant positive relationship was obtained between musical vocal skills and nonmusical speech production skills, implying that communication through music bypasses speech and language barriers in individuals with autism. Other than improvisational music therapy, the use of rhythmic patterns and melodies has been found to be beneficial for aiding learning and memory in autism. Buday (1995) observed that the number of signed and spoken words correctly imitated by children with autism was higher using sung rather than spoken text.

Music as a mode of intervention has yielded beneficial effects for the modification of behavioral problems. Burleson, Center, and Reeves (1989) obtained results to show that background music reduced off-task behaviors and facilitated task related performance by children. Orr, Myles, and Carlson (1998) investigated the effect of rhythmic entrainment on erratic classroom behaviors in their case study of a girl with autism. Rhythmic entrainment involves the use of music to aid relaxation by the introduction of externally produced rhythms, designed specifically to re-entrain the body to its natural rhythmic patterns. They found that rhythmic entrainment helped in reducing problematic classroom behavior. Focusing on the role of synchronization, Wimpory, Chadwick, and Nash (1995) used musical interaction therapy in order to synchronize music to interactions between a 3 year old girl with autism and her mother. The girl showed improvements in eye contact and other forms of social acknowledgement, and she initiated more episodes of interactive involvement with her mother. Moreover, a two-year follow up showed that these benefits were sustained.

Social stories are also often employed as a method of choice for individuals with autism. These short stories are designed to describe social situations extensively so as to provide relevant social cues that children with autism often fail to pick up in their natural environments (Kuoch & Mirenda, 2003). The aim is to improve awareness and understanding of various social situations and the appropriate responses to be elicited in such complex situations. Brownell (2002) presented social story information either musically or by simply reading the story contents to four first- and second-grade students with the aim of modifying behaviors. The results showed that musically presenting social stories was a better treatment alternative.

The above mentioned studies show that music therapy is an effective treatment choice for improving communicative abilities in autistic individuals. This improvement in communication has been brought about by attempts to improve emotion recognition, eye contact, behavior modification and so on. Even though music therapy has been used with success, there is still a shortage of theoretical papers that deal with understanding the neural substrates and cognitive mechanisms underlying the improvement in such adaptive skills. Towards this end, the current paper will discuss recent findings on the neuropsychology and neural bases of music to help make the picture clearer. The focus here will be on understanding why music therapy leads to an improvement of emotion recognition as part of communicative skills.

A COGNITIVE MODEL OF MUSIC THERAPY FOR ASD

Peretz (2001a, 2001b; Peretz & Coltheart, 2003), Dowling (2001) and Lerdahl and Jackendoff (1983) have proposed domain specific mechanisms and specialized neural networks devoted to music processing components principally involving pitch based computations (for instance, tonal encoding of pitch). Peretz & Coltheart (2003) have emphasized the functional modular nature for tonal encoding of pitch in music. This module may minimally overlap with language according to Peretz (2006), which is consistent with a previous finding that pitch variations generate a determinate scale in music but not in speech intonation contours (Balzano, 1982). Peretz (2006) further proposed that this particular module in music processing may be innately constrained. The module for tonal encoding of pitch is also hypothesized to be responsible for the perception of emotions in music (Peretz, 2002). Supporting this claim, Khalfa and colleagues (2008) showed that tonal variations of pitch support distinctions between happy and sad musical emotions. They used psychophysiological and psychometric indices of emotions to investigate the proposed hypothesis. Musical excerpts were created either by removing pitch variations (the rhythmic version) or by removing both pitch and temporal variations (the beat alone condition). The results showed that reliable happy/sad distinctions require tonal variations; the rhythmic version or tempo alone conditions did not support this emotional distinction. The authors further reasoned that tempo entrainment might exist in tempo alone conditions, but the effects might disappear when tempo is embedded within a musical context or with rhythm. The results of this study
are consistent with various other findings that show an advantage for decoding happiness and sadness over fear and anger in musical stimuli (Bunt & Pavlicevic, 2001; Terwogt & Van Grinsven, 1991).

Initially, it was hypothesized that since there were emotion processing deficits in autism, an emotionally responsive system for music might also fail to develop for individuals with the disorder (Peretz, 2001c). However, recent developments in cognitive science have changed the initial theorizing about the emotionally responsive system for music in autism. Individuals with autism are shown to be good at detecting pitch from musical tones (Mottron, Peretz & Me´nard, 2000). As previously stated, individuals with autism also show no deficits in processing affect in music (Heaton, Hermelin & Pring, 1999). Since Peretz (2002) hypothesized that the module for tonal encoding of pitch is responsible for the perception of emotions in music, it is likely that the ability of individuals with autism to recognize musical affect is based on the module for tonal encoding of pitch. Although scientific efforts to confirm this hypothesis have not been initiated, the claim is consistent with evidence that there are separate emotion recognition modules (Chuang & Wu, 2004), including neuroscientific evidence (Hornak, Rolls & Wade, 1996). Thus, music can be employed to alleviate emotion recognition deficits in individuals with autism because tonal pitch differentiation is the primary means by which emotions are perceived.

Some theoretical positions might argue that music therapy works effectively for individuals with autism due to other benefits that the therapy might confer such as increased attention from the class teacher, opportunities to use musical instruments (Kern and Aldrige, 2006) or the structure and predictability of therapeutic sessions (Wigram & Gold, 2006) [2]. Therefore, it is important to emphasize here that tonal encoding of pitch in musical sessions for autism might have different advantages. Firstly, Peretz (2002) argues that the synchronization involved in musical activity is an effective way to promote simultaneous dancing and singing (Brown, 2000). Synchronization in turn targets mirror neurons (Berrol, 2006) that are also found to be functioning atypically in autism (Williams, Whiten, Suddendorf & Perrett, 2001). Thus, there might be a different mechanism through which the tonal variations in pitch bring about effective results. Regardless, the main aim of the current article is to emphasize that the emotion recognition/distinctions (happy and sad) might be dependent upon sensitivity to tonal variation.

Along similar lines, Gagnon & Peretz (2000) showed that tonality could be put to different uses depending upon the task at hand. They obtained results to show dissociation between affective and non-affective judgments. Their experimental participants showed no hemispheric differences while judging melodies to be correct or incorrect (non-affective task). On the other hand, when required to judge the melodies as pleasant or non-pleasant (affective task), the participants did show hemispheric differences, where pleasant responses were indicative of a left hemisphere predominance and vice versa. Therefore, the use of tonality is made differently for affective or non-affective purposes. However, this result is not consistent with the results obtained by Bryden et al. (1982), which did not reveal hemispheric differences for positive and negative judgments. It is important to note that the study by Bryden and colleagues involved the manipulation of minor and major modes to yield pleasant or non-pleasant melodies, whereas the study by Gagnon & Peretz (2000) manipulated tonality. Therefore, it is quite possible that tonality and other musical cues that are responsible for emotional identification involve different mechanisms. Understanding the nature of such mechanisms is an important challenge for future research.

Future research efforts will yield valuable insights if geared towards investigating the mechanisms through which music therapy brings about desirable results in autism, especially if the fundamental goal is to improve the recognition of basic emotions such as happiness or sadness. The current literature identifies tonal variations in pitch as a potential mechanism to investigate. Cognitive and behavioral studies could be combined with brain imaging tools to understand the brain basis of such mechanisms, as well as to throw light on the neural processes that undergo changes following the application of musical intervention.

**Formulating a Rehabilitation Program Based on Music**

Understanding the cognitive and brain basis of the tonal encoding hypothesis will not be enough for planning rehabilitation programs unless this effort is combined with the aim of generalizing the results obtained in the laboratory. But what would be the basis of such generalization? As stated above, Peretz (2006) hypothesized that the module for tonal encoding of pitch in music may minimally overlap with language. If this is the case, then can we expect a generalization of effective music lessons to language in day-to-day functioning? Given that individuals with autism fail to interpret and recognize vocal
expressions of emotions, a reasonable goal of any intervention would be to improve recognition of emotions in daily functioning. Then how successful can an intervention program be if it is based on a module that minimally overlaps with language?

A look at the brain imaging literature may shed light on this issue. Patel (2003) posits significant overlap between modules for language and music processing and even points to the overlap in neural structures supporting language and music processing as evidence for the domain nonspecificity of music and language (Koelsch et al., 2002; Maess et al., 2001). For example, BA 47 has been shown to be associated with the processing of both linguistic and musical structure (Levitin and Menon, 2003), while BA 44 (Broca’s area) and its right hemisphere homologue have been found to be associated with the processing of musical syntax (Maess et al., 2001). But it is important to emphasize here that the existence of overlap between language and music for syntactic processing need not imply the non-domain specificity of other specific modules in music processing. Consistent with this claim, Peretz & Coltheart (2003) reviewed an imaging study that showed that regions such as the middle temporal gyrus, occipital lobe, and parietal lobe were activated during music perception. These regions are not typically activated in language processing tasks, supporting the claim that music processing involves various modules, some of which are domain specific for music (Schmithorst, 2005). It should be noted, however, that the significance of these activated neural regions during music perception is not clear.

At this stage, it might not be possible to discuss generalizations by considering overlaps between language and music. Generalizations might become more feasible if prosody was brought into the picture. Speech prosody refers to the musical aspects of speech that are used to convey emotions (Juslin & Laukka, 2003) and is used in a similar manner across cultures to convey emotions (Bolinger, 1978). Therefore, verbal comprehension or semantics are not necessary to understand emotions conveyed through speech, thereby facilitating the decoding of emotions even in a foreign language (Thompson, Schellenberg & Husain, 2004). Even evolutionary theorizing suggests that music, speech, prosody and facial expressions share common ancestry as spatial-temporal patterns (Dissanayake, 2000) perhaps laying the common ground for a single intervention attempt with favorable effects for alleviation of emotional processing deficits from speech, prosody and facial expressions.

Pitch and rhythmic information are identified as critical dimensions in both music and prosody (Patel et al., 1998). Both pitch and rhythmic cues retain their identities in musical tunes across various manipulations of pitch levels and tempo. In speech, pitch variations provide an important source of emotional information consistent with the results of Khalifa and colleagues (described above). Therefore, if music lessons help in decoding emotions from speech prosody then music interventions could greatly aid in adaptive living.

Thompson, Schellenberg & Husain (2004) investigated the effect of music lessons on the recognition of emotions from speech prosody. In their study, 6 year old children underwent musical (keyboard) training for a period of 1 year. After the end of the training period, these children were better at identifying happy and sad emotions from speech than children who underwent drama, singing or no art lessons at all for the same time period. The music group was better at the happy-sad distinction as compared to the fearful-angry distinction. These results imply striking parallels to the theorizing developed in the present paper. Firstly, the keyboard group was better at the identification of happy and sad emotions; these two emotions are hypothesized to be dependent upon the tonal encoding of pitch. Secondly, the keyboard group performed better than the groups taking other forms of training. These other types of training like drama and singing could also be considered forms of music therapy but are not as effective as keyboard lessons. Therefore, it is likely that tonal encoding of pitch underlies the ability to distinguish between happiness and sadness in music. Another noteworthy result of the study was the equivalent performance by keyboard and drama groups on fearful-angry comparisons. According to Thompson et al. this may occur because the drama lessons typically involve speech training and the use of prosody. Thus, it might also be possible that different sorts of music therapy could be employed for alleviating deficits for different emotional categories, and the tonal variations of pitch might be a better choice of intervention for focusing on the basic category of happy-sad expressions; a question that can only be answered with further investigations.

What remains to be seen with future studies on autism is whether music lessons in autism aids in recognition of emotions from speech prosody for one’s own and foreign language. If taking music lessons enhances the recognition of emotions by bypassing the language barriers then this will provide stronger support for the proposal outlined here. Thus, to comprehensively understand the nature of music therapy and its effectiveness for autism it is important to focus well coordinated research efforts
in two distinct directions. Research in the cognitive domain should progress with parallel efforts to evaluate whether the findings of such research can be generalized to various domains of daily functioning. Unless such endeavors are well planned and coordinated, attempts at applying the findings of basic research studies will not be beneficial. This holistic approach will not only improve our understanding of music therapy and musical emotions, but will also yield positive results for planning rehabilitation programs for children afflicted by autism. Therefore, the need of the hour is greater cooperation between researchers, parents, policy makers and the community at large. It is hoped that this manuscript will be a source of inspiration towards this direction.

CONCLUSION

The current article describes the cognitive basis of emotion recognition through music. According to the model developed here, tonal encoding of pitch provides the basis for recognition of emotions. This hypothesis explains why individuals with autism lacking the ability to perceive and interpret affective cues within the social domain are still able to process affect from music. Since individuals with autism are able to recognize emotions from music but are unable to process emotional information from voice and faces, it is worth mentioning that a global emotionally responsive system might be possible, consistent with the findings that there can be separate emotion recognition modules. The tonal encoding of pitch is further hypothesized to be domain specific and functionally modular in nature, even though specialized neural networks supporting it have not yet been explored. The present article lays emphasis on the need to build a rehabilitation program for promoting recognition of emotions in autism based on intact music modules, and also helps in elucidating the basis of generalization of results to everyday living.

Several questions await the results of future research and until such answers are forthcoming it is difficult to come up with a thorough rehabilitation program based on music therapy. First, there have been few studies in autism that show the effectiveness of music therapy. Second, there is a general lack of understanding about the mechanisms in music that promote emotion recognition in autism. Future research efforts that focus on tonal encoding of pitch may help to address such challenges. It is quite possible that this basic module greatly helps in the recognition of happy and sad emotions in the musical context. It might not be very effective for promoting the recognition of other emotions and therefore other types of musical interventions might need to be included as therapy progresses. Tonal encoding of pitch might have several other advantages rather than just promoting recognition of emotions. There might be long term effects by the promotion of synchronization that targets the mirror neurons functioning sub-optimally in autism. Towards this end, it might be helpful to employ brain imaging tools to study the effect of music therapy on the mirror neurons in the brain. But we also need to combine it with strong cognitive theorizing.

The next fruitful direction will be to understand the basis of generalizing the effective results obtained in the therapy sessions within the lab to everyday functioning. One likely candidate in this regard is to understand the effect of music therapy on the recognition of emotions from speech prosody. Studies with children have shown that music sessions are effective at promoting recognition of happiness and sadness as compared to other emotions from speech prosody. The future studies should concentrate on understanding the mechanisms of such generalizations and it might be likely that, again, the tonal variations of pitch are responsible for such effects. If such effects are obtained with prosody, and music lessons are shown to aid in the recognition of happy and sad emotions from prosody in the context of foreign languages for autism, we could be more confident about this potential mechanism. Once generalizations for prosody are understood, other elements could be brought in to understand how benefits obtained in the laboratory could be generalized to facial expressions, for example. Here again we need combinations of cognitive theories with brain imaging tools to help us understand such interplays. As music treatment with autism progresses, additional music tools such as group singing could be incorporated to target the improvement in the recognition of other emotions and the basis of further generalizations.

NOTES

[1] Email: nehakhetrapal@gmail.com
[2] I thank the editor for bringing up this point.
REFERENCES


