Proceedings of the 25th Anniversary Conference of the European Society for the Cognitive Sciences of Music, 31 July-4 August 2017, Ghent, Belgium
Van Dyck, E. (Editor)

Investigating the Development of Joint Attentional Skills in Early Ontogeny through Musical Joint Action

Kevin Kaiser[1], Marvin Heimerich[2]

Department of Systematic Musicology, University of Cologne, Germany
[1]kevin299@web.de, [2]mheimeri@smail.uni-koeln.de

ABSTRACT

Joint attention is an important basis for social interaction such as making music together. The emergence of joint attention in infants around the age of 9 to 15 months is studied through joint attentional skills, i.e. communicative, non-verbal, observable behaviors. Although there are plenty of studies investigating this period of emergence, research extending to a possible further enhancement of joint attentional skills alongside the infants' social cognitive development is rare. Therefore, the goal of our current study is investigating whether joint attentional skills show enhancement in the later course of development of typically developed children and thus potentially facilitate increasingly more complex social interaction. In the current paper, we first introduce the design of our observational study to investigate joint attentional skills in a natural, musical joint action setting. Secondly, we present our coding scheme, which comprises two categories social gaze and musical gestures. Concerning social gaze, we then discuss the code gaze following in comparison with that used in Carpenter et al. (1998). Our results suggest that our coding scheme is appropriate for our observational study and reflects the natural, interactive setting.

I. BACKGROUND

Almost all of us are familiar with the feeling to get involved in social interaction through music and one can easily see that music provides a variety of ways to engage in social interaction in our everyday life. For example, music may make you dance together with your friends, you could share an aesthetic experience with your partner, or people may play music together in an ensemble. Music making requires coordinated motor and social actions (Overy & Molnar-Szakacs, 2009) and, thus, can be considered as joint action, i.e. any form of social interaction in which the agents have to coordinate their actions to bring about a change in the environment (Knoblich & Sebanz, 2008).

Indeed, the importance of social components over complex acoustic features in music are pointed out by several authors repeatedly (e.g., Overy & Molnar-Szakacs, 2009; Kirschner & Tomasello, 2009, 2010; Cross, 2012; Aucouturier & Canonne, 2017). However, few (but, for example, Kirschner & Tomasello, 2009, 2010) investigate the basic social communicative capacities that are needed to engage in a joint musical activity, such as joint attention.

Joint attention is a set of non-verbal, communicative behaviors ("joint attentional skills") which develop during the so-called 9-month revolution around 9 to 15 months of age (Carpenter et al, 1998). These behaviors "indicate a newly emerging understanding of other persons as intentional beings whose attention to outside objects may be shared, followed into, and directed in various ways" (Carpenter et al., 1998: v). That is, joint attention is viewed as gradually arising social cognitive and interactive skills.

Joint attentional skills are observable by definition and include, for example, joint engagement, which is necessary to share attention with other social agents, and gaze or point following, which are needed to coordinate one's own attention with others in a social setting. Joint attentional skills are proposed to emerge in the following order, reflecting their increasing complexity: *sharing attention* which includes joint engagement, *following attention* which includes gaze and point following as well as imitative learning, and *directing attention* which includes declarative and imperative gestures to direct attention (Carpenter et al., 1998).

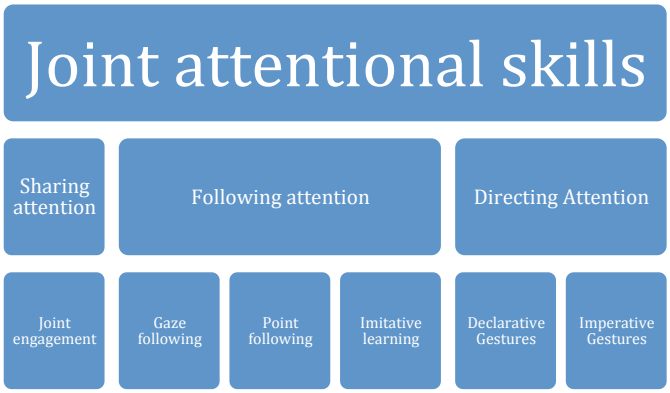

Figure 1. Chart depicting the joint attentional skills

The emergence of joint attention is linked to the emergence of the capacity to view other people as intentional agents (Tomasello, 1999) or, in other words, to recognize that others are 'like me', i.e. so-called *'like-me'-stance* (Meltzoff, 2007). This ability marks the beginning of intersubjective joint attentional interactions and joint attention can thus be regarded as a milestone of the social cognitive development.

Further social cognitive mechanisms such as *Empathy* and *Theory of Mind* develop in the later developmental stage. *Empathy* is a complex, high-level cognitive ability to simulate another person's mental as well as affective state (while being able to clearly differentiate between one's own and other's mental state) (Coplan, 2011). *Theory of Mind* is the ability to infer on one's own or others' mental states and reflect upon it while being able to clearly differentiate between self and other mental states. The former develops around the age of 1.5 to 2 and the latter emerges around the age of 4 (Bischof, 2009).

Because joint attention marks the beginning of complex social interactions through viewing others as intentional agents, it is possible that joint attention scaffolds those higher social cognitive mechanisms. Indeed, joint attentional skills have been suggested as "precursors to the understanding of the thoughts and beliefs of others that emerges at around 4-5 years of age" (Carpenter et al., 1998: 1), i.e. Theory of Mind.

Until now, the emergence of joint attention has been extensively studied, but research on a possible further

enhancement of joint attentional skills alongside the infants' social cognitive development is rare. Especially, research on a possible enhancement exceeding the period of the development of social cognitive mechanisms like Theory of Mind is still missing.

II. AIMS

The goal of our current study is investigating whether joint attentional skills show enhancement after the age of emergence, which is between 9 and 15 months of typically developed children.

Through linking a possible enhancement of joint attentional skills to joint musical interaction behavior, this study also aims at investigating a potential facilitation of enhanced joint attention to increasingly more complex social interaction. Especially, in the current paper, we present a coding scheme for analyzing joint attentional skills in a more natural, group interaction setting. Additionally, this coding scheme is discussed in comparison to the coding scheme of Carpenter et al. (1998).

III. METHOD

In order to investigate the possible enhancement of joint attentional skills an observational study was conducted with children of different age groups (see Section B: Participants) in a musical joint action setting. Music, because of its social, non-verbal communicative aspects, is an appropriate domain to study joint attention independently of children's language skills.

A. Structured Observation

In contrast to conventional experiments which require controlled environment, stimuli, and tasks, an observational study allows us to examine children's nonverbal behavior in a natural, less controlled environment as well as it enables us to investigate processes, i.e. behavior that unfolds in time (Bakeman & Quera, 2012). Observational studies can be structured or unstructured. Structured observations, which are used in the current study, include a predefined catalog of behaviors that is going to be observed (so-called "coding scheme"), as well as trained, unbiased observers, that code these behaviors (so-called "coders"). Unstructured observations are similar to narrative reports in that the observer decides which behavior is being observed (Bakeman & Gottman, 1997).

Coding schemes determine which behavior is being observed through a set of descriptive codes which represent certain type of observable behavior and are usually designed for a particular study according to the hypothesis being tested (Bakeman & Gottman, 1997). Coders analyze the behavior on the basis of the coding scheme either live or through reviewing previously video recorded materials (Bakeman & Gottman, 1997). In the current study, the latter method was used.

Because observational studies are faced with several problems regarding observer agreement like perception and memory errors as well as coder bias, observer agreement is necessary to obtain reliable data. Methods that are being used to ensure high observer agreement include training multiple coders and checking for agreement between coders using Cohen's Kappa (Bakeman & Gottman, 1997). Cohen's Kappa is a type of agreement statistic that corrects for two observers agreeing by chance (Bakeman & Gottman, 1997).

B. Current Study

1) *Participants*. Twenty-six children participated in the study which were grouped based on similar age as follows: Group 1 consists of 4 children, the age ranging from 1.5 to 2 y.o.; Group 2 consists of 6 children, the age ranging from 1.5 to 3 y.o.; Group 3 consists of 4 children, the age ranging from 3 to 3.5 y.o.; Group 4 consists of 4 children, the age ranging from 4 to 4-5 y.o.; Group 5 consists of 6 children, ranging from 5 to 6 y.o..

Table 1. List of Participants

Group	No. of children	Age range
1	4	1.5 - 2
2	6	1.5 - 3
3	4	3 - 3.5
4	4	4 - 4.5
5	6	5 - 6

2) *Session*. Every Group participated in one session of lessons of music education for young children. All sessions took place in the „Fröbel-Kindergarten Regenbogen" (Cologne, Germany) and were video-recorded with four cameras, placing one camera in each corner of the room (see figure below).

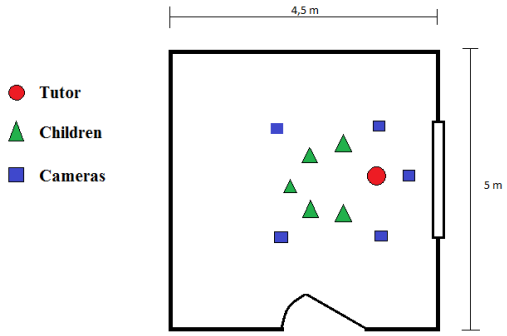

Figure 2. Layout of the room including the position of the tutor, the children, and the cameras.

The sessions were identical for every group and lasted about 30 minutes. The lessons consist of music making in a group setting under the guidance of a tutor and are designed for children to obtain rudimentary music skills and experience joy in making music and singing together. A lesson comprised six songs that were being played in succession.

The songs were interactive in nature and varying in tasks as well as instruments used. Each song included different "tasks": children imitated certain arm or leg movements, stomped or clapped in rhythm with the song and tutor and made certain facial and vocal expressions. The tutor played the guitar in all but two songs. In these two songs, the tutor and the children played either drums or claves together.

3) *Materials for analysis*. For the analysis, a short opening song was selected. We chose this song, as it provides less restrictive tasks (as in quantity and duration of tasks) as other

songs of the lessons and thus give children more freedom to interact with other subjects, the tutor or the group as they see fit. This will result in more natural and spontaneous intersubjective behavior.

This song is performed three times in total: First, the tutor is singing the song with the children (a "welcome" song) in a normal manner. Afterwards the "task" is to sing this song as quietly as possible. The song concludes with singing the song as loud as possible. At the end of each of the three songs, the children pound on the floor in a "normal", quiet and loud manner, respectively. The duration of the "set" (i.e. the song sung three times) is about 1.5 minutes in total.

IV. RESULTS

A. Coding Scheme

To analyze the abovementioned joint attentional skills and non-verbal interactive behavior of children across a broad age range, we developed a coding scheme, which includes two categories: social gaze and musical gestures.

Social gaze is comprised of two code events: gaze and gaze following. The former is described as the act of switching one's gaze to a target, while the latter as the act of following one's gaze to a target. Gaze following was chosen as a code, because it is a key component of joint attention (Falck-Ytteret et al., 2015) and easily observable in a natural interactive setting. By investigating the frequency of gaze in relation to gaze following, we are able to examine a possible enhancement of gaze following and thus joint attention.

Musical gestures include musical joint action behavior such as rocking, singing and clapping together and enable us to investigate how and how intensive children interact with others within our musical joint action setting.

Moreover, by examining the relation between those two categories, it is possible to investigate whether children's interactive joint action behaviors are facilitated by enhanced joint attentional skills.

Coders are tasked to record behavior on the basis of this coding scheme. The code gaze following, a central code to our study, is described as follows: when a subject (A) changes his gaze from one subject (B) to a second subject (C) by following the gaze of the first subject (B) to the second subject (C).

In contrast, the study of Carpenter et al. (1998) used a more elaborated description of the code gaze following. For example, the initiation and process of the gaze following task was under the experimenter's control:

1) Initiation
   a) The infant plays with an uninteresting toy to begin with;
   b) Experimenter 1 (E1) calls the infant by name and waits for him to create eye contact before looking at the assigned target.
2) Process:
   a) After establishing eye contact, E1 looks at the assigned target with "an excited facial expression and vocalization (a gasp)" (Carpenter et al., 1998: 39);
   b) E1 alternates her gaze between the infant's eyes and the target several times waiting for the infant to respond with following her gaze.

B. Comparison of the Coding Scheme

The code gaze following is described differently in Carpenter et al. (1998) and our current study, which reflects the unique setting of each study.

Carpenter et al. (1998) are utilizing a more controlled setting and thus create an "unnatural" situation in which the infant is expected to perform gaze following. This situation is created through the experimenter calling the infant by its name and waiting for eye contact, thus controlling for the initiation of gaze following.

The code description for "gaze following" in our study is reflecting the more naturalistic design of our study, in which the children interact much more freely with one another. Thus, our code description leaves great freedom regarding for example the initiation of gaze following.

The differences between the description of the code "gaze following" in Carpenter et al. (1998) reflect the differentiation between controlled and naturalistic measurement methods regarding gaze following discussed by (Falck-Ytteret et al., 2015). By not controlling the setting and external factors as much as Carpenter et al. (1998) and not being able to control the children's behavior, our codes are much less restrictive, but also, more natural and better suited for our study.

V. CONCLUSION

To the best of our knowledge, our study is the first empirical developmental study that studies joint attentional skills and children's interactive behavior by analyzing gaze and musical gestures in an ecologically valid musical joint action setting in such a broad age range. In the current paper, we presented and discussed the design of the coding scheme in detail.

By coding social gaze, i.e. gaze and gaze following as well as analyzing between-group differences in frequencies of gaze and gaze following, we are able to study a possible enhancement of joint attentional skills. Musical gestures allow us to determine possible changes in children's interactive behaviors. By examining the relationship between gaze and musical gestures, it is possible to investigate whether children's interactive behaviors are facilitated by enhanced joint attentional skills.

By comparing our coding scheme, especially our code description of gaze following to that of Carpenter et al (1998), we showed that both our coding scheme and the coding scheme of Carpenter et al. (1998) reflect the design of each setting: our natural, observational study stands in contrast to the more controlled setting used by Carpenter et al. (1998).

ACKNOWLEDGMENT

First and foremost we would like to thank our supervisor Rie Asano for the continued support of our research and revisions of our written abstracts and papers. Secondly, we would like to thank Lena Irnich for allowing us to conduct this study as well as Sabrina Kierdorf, Kimberly Severijns, Bela Hardekopf and Sascha Bullert for contributing to this research

in a variety of ways. Lastly, we would like to thank Julia Feltes for testing our codes with us.

REFERENCES

Aucouturier, J. J., & Canonne, C. (2017). Musical friends and foes: The social cognition of affiliation and control in improvised interactions. *Cognition, 161*, 94-108.

Bakeman, R., & Gottman, J. M. (1997). *Observing Interaction: An Introduction to Sequential Analysis* (2nd ed.). New York, NY: Cambridge University Press.

Bakeman, R., & Quera, V. (2012). Behavioral observation. In H. Cooper (Ed.-in-Chief), P. Camic, D. Long, A. Panter, D. Rindskopf, & K. J. Sher (Assoc. Eds.), *APA Handbooks in Psychology: Vol. 1. APA Handbook of Research Methods in Psychology: Psychological Research, Foundations, Planning, Methods, and Psychometrics*. Washington, DC: American Psychological Association.

Bischof, N. (2008). *Psychologie: ein Grundkurs für Anspruchsvolle,* Stuttgart: Kohlhammer.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, M., Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development, 63*(4), 1-174.

Coplan, A. (2011). Understanding empathy. In *Empathy: Philosophical and Psychological Perspectives*. New York, NY: Oxford University Press.

Cross, I. (2012). Music as a social and cognitive process. In P. Rebuschat, M. Rohrmeier, J. A. Hawkins, I. Cross (Eds.), *Language and Music as Cognitive Systems*. New York, NY: Oxford University Press.

Falck-Ytter, T., Thorup, E., & Bolte, S. (2015). Brief report: Lack of processing bias for the objects other people attend to in 3-year-olds with autism. *Journal of Autism and Developmental Disorders, 45*(6), 1897-1904.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102(*3), 299-314.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes proscocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(3), 354-364.

Knoblich, G., Sebanz, N. (2008). Evolving intentions for social interaction: From entrainment to joint action. *Philosophical Transactions of Royal Society B, 12*(36), 2021-2031.

Overy, K & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. *Music Perception, 26*(5), 489-504.