Robot-Assisted Training of Joint Attention Skills in Children Diagnosed with Autism

Jasmin Kajopoulos[1], Alvin Hong Yee Wong[2], Anthony Wong Chen Yuen[2], Tran Anh Dung[2], Tan Yeow Kee[2], and Agnieszka Wykowska[3,4](✉)

[1] Neuro-Cognitive Psychology Master Program, Department of Psychology, Ludwig-Maximilians-Universität, Munich, Germany
jasmin.kajopoulos@campus.lmu.de
[2] Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore
{hyawong,cywong,tanhdung,yktan}@i2r.a-star.edu.sg
[3] General and Experimental Psychology Unit, Department of Psychology, Ludwig-Maximilians-Universität, München, Germany
agnieszka.wykowska@psy.lmu.de
[4] Chair for Cognitive Systems, Technische Universität München, Munich, Germany

Abstract. Due to technological and scientific advances, a new approach to autism therapy has emerged, namely robot-assisted therapy. However, as of now, no systematic studies have examined the specific cognitive mechanisms that are affected by robot-assisted training. This study used knowledge and methodology of experimental psychology to design a training protocol involving a pet robot CuDDler (A*STAR Singapore), which targeted at the specific cognitive mechanism of responding to joint attention (RJA). The training protocol used a modified attention cueing paradigm, where head direction of the robot cued children's spatial attention to a stimulus presented on one of the sides of the robot. The children were engaged in a game that could be completed only through following the head direction of the robot. Over several weeks of training, children learned to follow the head movement of the robot and thus trained their RJA skills. Results showed improvement in RJA skills post training, relative to a pre-training test. Importantly, the RJA skills were transferred from interaction with the robot to interaction with the human experimenter. This shows that with the use of objective measures and protocols grounded in methods of experimental psychology, it is possible to design efficient training of specific social cognitive mechanisms, which are the basis for more complex social skills.

Keywords: Autism · Robot-assisted therapy · Joint attention · Social robotics

1 Introduction

1.1 Autism

As medical and neurobiological knowledge advance, early detection tools and intervention methods for Autism Spectrum Disorder (ASD) have been growing [1,2]. Deficits in joint attention skills have been found to be one of the earliest signs for

A. Tapus et al. (Eds.): ICSR 2015, LNAI 9388, pp. 296–305, 2015.
DOI: 10.1007/978-3-319-25554-5_30

ASD, wherein impairments in joint attention (i.e., attending to where others attend with the awareness that the others attend to the same location/object) were associated with stunted language development and decreased social and communicative skills [3,4,5]. Thus, research into joint attention, especially the development of intervention methods, as well as understanding the mechanisms behind joint attention has become a pivotal area of interest in autism research [6].

1.2 Robot Therapy

In recent years with rising technological advances, especially in the field of robotics, a new area of intervention method for children with ASD has been emerging, namely social robots as a tool for autism therapy [7-13]. The therapeutical value of social robots is grounded in reliability, simplicity, and predictability of a robot's behavior [7,12]. However, although in recent years various social robots for autism therapy have been constructed [7,11,12], to date there is still a limited number of training protocols which would use social robots and target at specific neuro-cognitive mechanisms. For example, Warren and colleagues [14] showed that after 4 sessions with a robot, children seemed to improve in following joint attention cues given by the robot. However, it was not demonstrated to what extent this effect prevailed in other situations, i.e. whether also general joint attention skills of the children (independent from the robot) improved. Similarly, it was shown that other prosocial behaviour, such as interaction through gaze and touch, was elicited and improved through the use of robots [15]. But again testing of the effect seemed to have been limited to the robot. In general, autism robot therapy research could benefit from independent diagnostic measurement [16]. On the other hand, Wainer and colleagues [17] demonstrated through interviews and questionnaires that improvements in interaction skills elicited by the deployment of an after-school robotics class seemed to persist even after the class. Thus, there is some evidence that social interaction skills, such as joint attention, may be improved through robotic therapy. However, the field is in need of systematic evaluations (with reliable objective methods of experimental psychology) of the effectiveness of robots' therapeutic use, independent of a specific robot platform.

1.3 Aim of Study

The present study aimed at designing a training protocol of robot-assisted therapy for autism that would extract and target at specific mechanisms of social cognition, and would also allow for experimental assessment of the efficiency of such protocol. To that end, we used a social robot CuDDler (designed by the Agency of Science and Technology (A*STAR) Singapore) and we embedded a variation of a spatial-attention cueing paradigm in a game that children played with the robot. Spatial attention of children was cued by the robot's head direction. Based on previous research in experimental psychology [3,18-20] we aimed at extracting and training the mechanism of responding to joint attention (RJA), which is the fundament for other social cognitive skills. Further, we aimed to test whether the trained joint attention skills prevail independently of the robot's presence by using the abridged Early Social Communication Scale (ESCS, [21]), which has been shown to be effective in measuring children's behaviour in naturalistic environments [22,23].

2 Materials and Methods

2.1 Participants

7 children (Mean age: 4.6, age range: 4 to 5 years, 4 male) all diagnosed with ASD (1 low, 4 low/moderate, 1 moderate, 1 moderate/high functioning) took part in the experiment. All children were right-handed and all had normal vision. Parents were recruited via the early intervention center THK EIPIC Centre (Singapore), and they volunteered their children for participation. All parents gave written informed consent regarding the participation of their children in the experiment. The experiments were conducted at the THK EIPIC Centre (Singapore) in collaboration with A*STAR Singapore and the Ludwig-Maximilians-Universität Munich. The experimental procedures consisted of purely behavioural data collection (e.g., accuracy rates), and were video recorded. The procedures did not include invasive or potentially dangerous methods and were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Data were stored and analyzed anonymously.

2.2 Stimuli and Apparatus

The key apparatus in the experiment was a robot CuDDler (A*STAR), controlled via a computer interface (Windows 7) and a smartphone (Google Nexus 4). The sequence of movements and speech were started by pressing different buttons on the interface. Additionally, picture stimuli were presented on two 136.6 x 70.6 mm mobile phone screens (bephone, 640 x 480), located to the left and right of the robot at a distance of ~ 40 cm (11° of participants' visual angle), tilted ~ 45° relative to the robot, so that the robot seemed to "see" stimuli on the screens after turning its head (cf. Fig. 1, left).

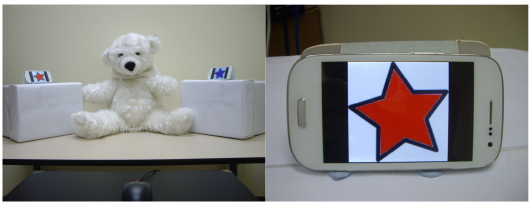

Fig. 1. Experimental Setup (left) and an example of a stimulus (right).

The target stimuli were 10 colorful line drawings of various objects (star, apple, ball, candle, flower, hat, heart, ice cream, plane, sweet, Fig.1, right) in 4 colors: red, blue, green or yellow. In one trial, there were always two drawings of the same object presented – one on each phone screen in different (randomly chosen) colors. In one set of 20 trials all objects in all colors appeared once. The stimuli were fit to the center of the phone screens (136.6 x 70.6 mm) and covered approximately 2° in height and approximately 3.5° in width of visual angle of participants. The robot's head moved pseudo-randomly to the left or right approximately 2.3° in participants' visual angle from the midline with a 50% chance occurrence of either side. The participants were seated 200 cm from the robot, behind a table with a mouse for giving the response.

2.3 Procedure

The protocol's task difficulty, in particular giving color responses as well as abilities to press buttons were first discussed with the children's teacher, who confirmed that in principle all children had learned to discriminate colors and to express this verbally. The whole training protocol consisted of 3 phases. First, a so-called "pre-test" was administered by a human experimenter to measure children's joint attention skills – both responding to joint attention (RJA) and initiating joint attention (IJA) – using the abridged Early Social Communication Scale (ESCS, [21]) before the robot training was conducted. Then, the robot training, consisting of 6 sessions of approx. 20 minutes each, was performed over a period of approx. 3 weeks. 2-3 days after the last robot session, a "post-test" was administered (ESCS) by a human experimenter to once again measure children's joint attention skills. The pre- and post-test lasted approx. 10 minutes each, during which the human experimenter and the child were sitting in front of each other across a table. As we were interested in measuring joint attention skills, only 3 parts of the ESCS were used, the Object Spectacle Task (1 x), the Gaze Following Task (2 x) and the Book Presentation Task (2 x). For the object spectacle task children were handed (6 times) a different toy to play with, and it was tested whether the child would initiate a joint attention bid by gaze or gestures (IJA). During the book presentation task, an examiner pointed to 6 different locations in a book to test whether the child would follow this joint attention bid with gaze or pointing (lower RJA, as the object was in closer spatial proximity) or whether the child would initiate a joint attention bid by gaze or pointing (IJA) to different parts of the book. In the gaze following task, the examiner pointed to 1 of 4 posters located behind, to the left and right of the child (90° off midline) to test whether the child would follow this joint attention bid (higher RJA, as the object was further away).

Robot Training. Before the experiment, children were taught by a teacher or an experimenter to associate the mouse buttons with the respective phone screens (left-to-left/right-to-right mapping). The children were then instructed to follow CuDDler's head movement, and their task was to verbally (out-loud) name the color of the target, at which the robot was looking, and additionally press the corresponding button. In practice trials, an experimenter or teacher also pointed to the correct screens to help understanding of instructions. This lasted approx. 10 – 20 trials depending on children's proficiency, understanding and attention span. 1 session of the experiment (subsequent to practice) typically consisted of 20 trials, although this might have varied depending on children's proficiency, understanding and attention span.

Trial Sequence. A trial (Fig. 2) started with an experimenter pressing a button on the interface, whereupon the robot's head started to move from the starting position (looking straight ahead) either to the left or to the right screen (~2 s) while saying: "Look! A 'object name' ". Then the stimuli appeared on both screens and the robot said "What color is this?" Afterwards, the children had to give a verbal color response and press the correct mouse button (left or right), Fig. 2.

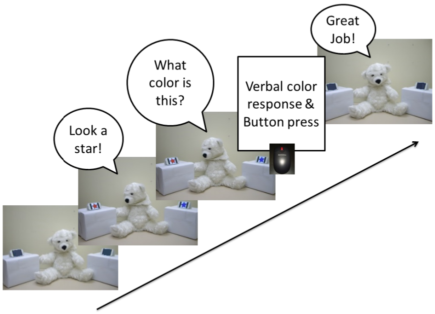

Fig. 2. Trial sequence: The sequence of events for one example trial.

When the children did not answer, the experimenter made the robot call the children's name, and repeat "What color is this?" and when the children still did not answer, the experimenter made the robot say "What color is the object I'm looking at?". As soon as the children pressed the button, the picture stimuli disappeared from both screens. Thus, the trial continued only when the children pressed a button. When the children only gave a verbal response, the robot was made to say "Please press the button". When the children still didn't respond, they were prompted by the instructor to press the button. After the response, the experimenter made the robot praise the child by saying "Good job!" and simultaneously making the robot move its head and arms "excitedly" in the direction of the child. Then the robot's head moved back to the starting position. In case a child gave only a button response, the robot said "Please tell me the color next time". In cases when a child did not respond, the teacher or instructor prompted the children to give a verbal response.

2.4 Data Analysis

Data of one child (low-functioning ASD) were excluded from all analyses, because they did not perform the required task. Also, data of two children were excluded from the analysis of RJA scores due to behavioral problems during the test, which led to a large difference in pretest and posttest in the RJA tasks, i.e. the number of opportunities for the examiner to start a joint attention bid in the RJA tasks drastically differed between pre- and posttest (average difference for the 2 excluded children = 6, average difference for the 4 children that were included in the analysis = 0.5, SD = 0.5).

Joint Attention Scores. To test whether children had higher joint attention scores after the robot training in comparison to before the robot training, the number of instances of RJA or IJA were calculated based on video recordings and according to the

ESCS coding sheet [21]. Initiating joint attention (IJA) scores were calculated based on the number of instances a child showed an IJA bid. The sum of every joint attention bid by the child amounted to the "total IJA" score. Therefore, higher score means more instances of IJA. Responding to joint attention scores (RJA) were based on percentage of times a child followed an examiner`s joint attention bid (pointing to a poster or to pictures in a book). The score was calculated separately for higher level (pointing to posters) and lower level RJA bids (pointing to a book). For lower level RJA (lRJA), a point was given when a child followed the examiner`s pointing gesture with their gaze. The sum of lRJA points was divided by the number of times the examiner visibly pointed to a picture in the book multiplied by 100 and thus the percentage score with a maximum of 100 (%) was calculated for lRJA. For the higher level RJA (hRJA), a score point was given when the child followed the examiner`s pointing gesture by sufficiently turning their head to indicate that they were looking at the poster. Subsequently, a paired t-test was administered to compare the average instances of joint attention (RJA or IJA) between pre- and post-test.

Color Accuracy. For these analyses, data from Session 1 counted as practice. Mean accuracy of color response for each child and each subsequent session was calculated. A paired t-test between Session 2 and 6 was administered to calculate the effect of the training over time. All trials, in which one of the following occurred were counted as errors and excluded: (1) the teachers pointed towards the screens to prompt a child, (2) a child didn't respond, (3) a child said a completely different color than what was displayed on the screen, or (4) a child responded with the wrong color first, but then named the correct color. However, if a child first responded with the correct color, but then said the other color, this was counted as a correct response. In cases when children named the red color as brown, the response was counted as correct. Similarly, when children said green instead of yellow and vice versa (only when yellow or green appeared with red or blue), the response was counted as correct, as the yellow and green color were similar. However, when both green and yellow appeared together, mixing up yellow and green was counted as an error, as the two simultaneously presented colors were easily distinguishable.

Button Press Accuracy. Also for these analyses, data from the first session counted as practice. Mean accuracy of button response for each child and for each subsequent session was calculated. A paired t-test between Session 2 and Session 6 was administered to calculate the effect of the training over the sessions. All trials, in which the teachers pointed towards the screens to prompt the children, were excluded.

3 Results

3.1 Joint Attention Scores

The analysis of joint attention scores showed that an average total RJA score of 193.75 (SD=12.5) in the posttest (after robot training) was significantly higher than the average total RJA score of 183.33 (SD=13.6) in the pretest (before robot training), $t(3) = 2.61$, $p = .040$, $d_z=0.797$ (cf. Fig. 3, left). When splitting the data into lower level RJA (lRJA) and higher level RJA (hRJA), lRJA results showed a significantly higher average lRJA score of 100 (SD=0) in the post-test, relative to an average lRJA score of 83.3 (SD=13.60) in the pre-test, $t(3) = 2.45$, $p = .046$, d_z=1.732 (cf. Fig. 3, right). The difference in the average hRJA score (pretest: 100, SD=0 vs. posttest: 93.75, SD=12.5) was not significant, $t(3) = 1$, $p = .80$ (cf. Fig. 3, right).

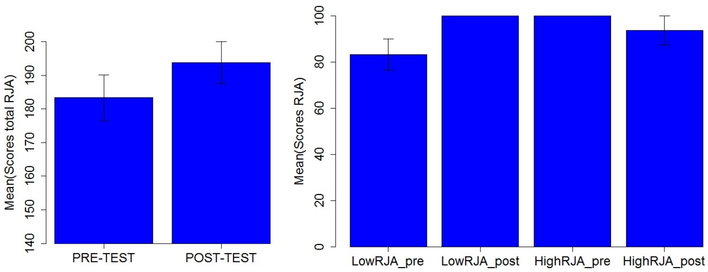

Fig. 3. Mean total RJA scores (left), lower RJA and higher RJA scores (right) in pre-test and post-test as measured by the ESCS [21] (left). Error bars: standard errors of the mean (SEM).

IJA scores were not significantly different in pretest (2.67, SD=2.42) vs. posttest (5.67, SD=5.27), $t(5) = 1.39$, $p = .112$.

3.2 Color Accuracy and Button Press Accuracy

Mean color accuracy in Session 6 (M= 0.69, SD= 0.46) increased significantly from Session 2 (M= 0.52, SD= 0.50), $t(5) = 2.19$, $p = .0.04$, cf. Fig. 4, left. Button press accuracy increased (marginally significant) for Session 6 (M= 0.75, SD= 0.43) vs. Session 2 (M= 0.67, SD= 0.48), $t(5) = 1.60$, $p = .086$, see Fig. 4, right.

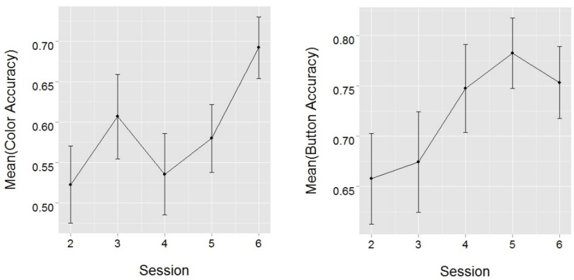

Fig. 4. Mean color (left) and button press accuracy (right) for Session 2 – 6. Error bars: SEM.

4 Discussion

Our study examined if a newly-developed robot-assisted training protocol embedding an attention-cueing paradigm in a game that children play with a robot has the potential of improving joint attention skills in children diagnosed with ASD.

4.1 RJA & IJA Scores

The present results showed a significant difference between RJA scores in pre- vs. post-training test, as measured by the ESCS [21], with greater scores after the robot-assisted training. The effect was large (as indicated by Cohen's d_Z effect size values), despite small sample. First, and foremost, this demonstrates that robot training was successful in increasing the children's skills in responding to joint attention, and that children transferred these skills to the interaction with human experimenter. Interestingly, when splitting the data into lower RJA (lRJA) and higher RJA (hRJA) scores, only lRJA scores showed a significant increase after the robot training. However, this may be simply due to the measurement's sensitivity not being high enough, as all children of this sample seemed to have a perfect score on hRJA skills already before the robot training. Possibly with a sample of children less skilled in hRJA before the training (or a more sensitive measure), a significant increase of hRJA skills might be observed after robot training. Furthermore, IJA scores were of equal magnitude before and after the training. This is in line with our hypothesis, as we designed a paradigm to specifically train RJA to maximize the effectiveness of the training. The sole effect of an increase in RJA skills and lack of influence of the robot training on the IJA skills supports the idea that there are two separate mechanisms underlying joint attention, namely IJA and RJA [3, 24]. Therefore, our results are in line with previous research and confirm that with carefully designed training protocols one can extract and target specific mechanisms of social cognition. Such focused training protocols should be effective as an intervention method in autism therapy [3, 24]. Moreover, the use of an independent diagnostic (ESCS), allows for suggesting that our findings are relatively robust and independent of the robotic platform.

4.2 Verbal Color Accuracy and Button Press Accuracy

Our present results indicated an increase in color accuracy over time and, in particular, a significant increase of accuracy between Session 2 and 6. This may be due to that children were increasingly more engaged in the game with the robot, arguing in favour of the training protocol. Additionally, a more general learning effect might have also played a role in the improved performance in color accuracy task. However, general learning is likely not to be the only factor influencing the effect, as it seems more plausible that the increase in performance during the robot-assisted training is related to the nature of the training protocol itself (completion of the task is possible only through following the robot's head direction). This is predicated on the fact that instances of RJA increased over the course of training, and this presumably had an effect on the increase of color accuracy. Similarly, results also indicate a general trend of increased button press accuracy over time. The fact that the differences between Session 2 and 6 were only marginally significant might have been related to that the button presses (and their mapping to respective stimuli) might have been demanding for the children with ASD, due to, for example, deficits in motor skills [6,25].

4.3 Implications and Future Directions

In general, results of this study provide convincing evidence for robot-assisted training of social skills in children diagnosed with ASD. The training can be effective when it is grounded in paradigms of experimental psychology that target specific cognitive mechanisms and measure improvement with objective measures. In our study, despite a relatively small sample size of children diagnosed with ASD, the robot CuDDler in combination with the attention-cueing paradigm proved appropriate to improve RJA skills. Improving RJA skills is crucial for social skills, as RJA is positively related to language development, social and communication skills of children diagnosed with ASD [3,5,26]. Thus, robot-assisted therapy which improves joint attention may – in a long run – also facilitate language learning and more complex social cognitive skills. Most importantly, to evaluate joint attention improvement, we used a diagnostic tool that was independent of the robot therapy, which suggests robustness of the effects. However, we assessed joint attention skills shortly after the last robot therapy session (post-test: 2-3 days later). Thus, future research should elucidate whether the improvement of RJA skills of children diagnosed with ASD after a robot-assisted therapy have a longer-lasting effect. Furthermore, follow-up studies (with larger sample sizes) should test if a longer training would be related to how permanent the improvement in joint attention skills is, and the degree with which it is transferred from human-robot to human-human interaction. Moreover, as some of the children did not want to participate in the tasks, future research will need to develop means to make the robot training more attractive (possibly a longer familiarization phase preceding the training) to children who are initially reluctant to participate.

4.4 Conclusions

In sum, this study is a first approach to design – based on paradigms from experimental psychology – training protocols that isolate and target specific cognitive skills of children with ASD. By combining information and knowledge gained from various interdisciplinary fields, such as robotics, cognitive neuroscience and psychology, one can design a training method that may lead to distinct and positive conclusions as to the effectiveness in improving joint attention skills in children with ASD. Future studies may take this as a basis on which to expand on, to examine this therapy method in more detail, test it on a larger sample size over a longer period of time and possibly also test variations of this protocol to train and improve other cognitive mechanisms.

Acknowledgment. We thank THK MORAL EIPIC for their help in recruiting participants and in preparation for the experimental study.

References

1. Matson, J.L., Kozlowski, A.M.: The increasing prevalence of autism spectrum disorders. Res. Autism Spect. Dis. **5**, 418–425 (2011)
2. Wing, L., Potter, D.: The epidemiology of autistic spectrum disorders: is the prevalence rising? Ment. Retard. Dev. D R **8**, 151–161 (2002)
3. Meindl, J.N., Cannella-Malone, H.I.: Initiating and responding to joint attention bids in children with autism: A review of the literature. Res. Dev. Disabil. **32**, 1441–1454 (2011)

4. Whalen, C., Schreibman, L.: Joint attention training for children with autism using behavior modification procedures. J. Child Psychol. Psyc. **44**, 456–468 (2003)
5. Mundy, P., Sigman, M., Kasari, C.: A longitudinal study of joint attention and language development in autistic children. J. Autism Dev. Disord. **20**, 115–128 (1990)
6. Charman, T.: Why is joint attention a pivotal skill in autism? Philos. T. Roy. Soc. B **358**, 315–324 (2003)
7. Cabibihan, J.J., et al.: Why Robots? A Survey on the Roles and Benefits of Social Robots in the Therapy of Children with Autism. Int. J. Soc. Robot. **5**, 593–618 (2013)
8. Dautenhahn, K.: Roles and functions of robots in human society: implications from research in autism therapy. Robotica **21**, 443–452 (2003)
9. Scassellati, B., Admoni, H., Matarić, M.: Robots for Use in Autism Research. Annu. Rev. Biomed. Eng. **14**, 275–294 (2012)
10. Anzalone, S.M., et al.: How children with autism spectrum disorder behave and explore the 4-dimensional (spatial 3D+ time) environment during a joint attention induction task with a robot. Res. Autism Spect. Dis. **8**, 814–826 (2014)
11. Bekele, E., et al.: Pilot clinical application of an adaptive robotic system for young children with autism. Autism **18**, 598–608 (2014)
12. Tapus, A., Mataric, M.J., Scasselati, B.: Socially assistive robotics [Grand challenges of robotics]. IEEE Robot. Autom. Mag. **14**, 35–42 (2007)
13. Tapus, A., et al.: Children with autism social engagement in interaction with Nao, an imitative robot–A series of single case experiments. Interact. Stud. **13**, 315–347 (2012)
14. Warren, Z.E., et al.: Can Robotic Interaction Improve Joint Attention Skills? J. Autism Dev. Disord. 1–9 (2013)
15. Robins, B., Dautenhahn, K., Te Boekhorst, R., Billard, A.: Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? Universal Access in the Information Society **4**, 105–120 (2005)
16. Diehl, J.J., Schmitt, L.M., Villano, M., Crowell, C.R.: The clinical use of robots for individuals with autism spectrum disorders: A critical review. Res. Autism Spect. Dis. **6**, 249–262 (2012)
17. Wainer, J., Ferrari, E., Dautenhahn, K., Robins, B.: The effectiveness of using a robotics class to foster collaboration among groups of children with autism in an exploratory study. Pers. Ubiquit. Comput. **14**, 445–455 (2010)
18. Friesen, C.K., Kingstone, A.: The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. Psychon. B Rev. **5**, 490–495 (1998)
19. Wiese, E., Wykowska, A., Zwickel, J., Muller, H.J.: I see what you mean: how attentional selection is shaped by ascribing intentions to others. PLoS One **7**(9), e45391 (2012)
20. Wykowska, A., Wiese, E., Prosser, A., Müller, H.J.: Beliefs about the Minds of Others Influence How We Process Sensory Information. PLoS One **9**(4), e94339 (2014)
21. Mundy, P., et al.: A Manual for the Abridged Early Social Communication Scale (ESCS), U.o. Miami (2003)
22. Noris, B., et al.: Investigating gaze of children with ASD in naturalistic settings. PLoS One **7**(9), e44144 (2012)
23. Noris, B., Keller, J.B., Billard, A.: A wearable gaze tracking system for children in unconstrained environments. Comput. Vis. Image Und. **115**, 476–486 (2011)
24. Taylor, B.A., Hoch, H.: Teaching children with autism to respond to and initiate bids for joint attention. J. Appl. Behav. Anal. **41**, 377–391 (2008)
25. Piek, J.P., Dyck, M.J.: Sensory-motor deficits in children with developmental coordination disorder, attention deficit hyperactivity disorder and autistic disorder. Hum. Movement Sci. **23**, 475–488 (2004)
26. Murray, D.S., et al.: The Relationship Between Joint Attention and Language in Children With Autism Spectrum Disorders. Focus on Autism and Other Developmental Disabilities **23**, 5–14 (2008)