Improving Visual Focus of Attention of Children with ASD during Prompted Task Execution using NAO Humanoid Robot

by

Di Xue

A thesis submitted in conformity with the requirements
for the degree of M.A.Sc
Graduate Department of IBBME
University of Toronto

Abstract

Improving Visual Focus of Attention of Children with ASD during Prompted Task Execution using NAO Humanoid Robot

Di Xue
M.A.Sc
Graduate Department of IBBME
University of Toronto
2015

Contents

Chapter 1

Research Objectives

1.1 Overall Goal and Approach

Our overall goal is to increase child with ASD's engagement level during COACH prompting and task execution, and thus improving prompt compliance and task completion rate. Our approach is:

1. to incorporate a half body humanoid robot, NAO T14 (see Figure 1.1) by Aldebaran Robotics, into the current COACH setup, capable of delivering verbal and gesture prompts and attention grabbers.

2. to automatically track the VFOA of child for more effective maintenance of child's attention. For example, by being able to recognize whether child is looking at the robot, the robot can call out child's name with a waving gesture or blink its LEDs for getting child's attention before prompting.

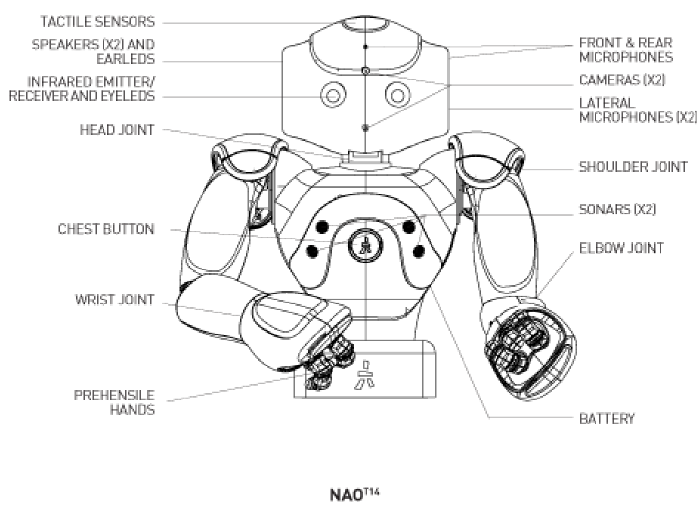


Figure 1.1: The Half-Body Version of Humanoid Robot NAO

1.2 Central Hypothesis

We hypothesize that the incorporation of an embodied agent, such as the humanoid NAO, is sufficient in better engaging child and better capturing and maintaining attention during prompting and task execution, and ultimately yields higher prompt compliance and task completion rates when assisting child with ASD through ADLs such as hand washing.

1.3 Specific Objectives

Our objectives are:

1. To investigate if a prompting system using NAO is able to guide child with ASD through hand-washing.

2. To explore the different modes of interactions between NAO and child when prompting hand-washing steps using a Wizard of Oz setup, focusing on verbal, gestures, and gaze for the modes of interactions

3. To implement a real-time algorithm for tracking child's VFOA

Our hypotheses are:

1. The humanoid robot, NAO, is able to independently assist child with ASD through hand-washing, and child exhibits greater engagement level, higher prompt compliance rate, and better task completion when prompted by NAO than by parent.

2. Gestural, gaze, and verbal are the essential modes of interactions present in the hand-washing prompting scenario between child with ASD and the prompting agent NAO.

3. Using 3DMM and ALR for estimating head pose and eye pose, and using the Kinect camera, a classification rate of more than 80% is achieved for estimating child's VFOA on NAO, monitor screen, soap, towel, tap region, hands, and idling.

Chapter 2

Wizard of Oz

One major objective of this thesis is to investigate the impacts that using a humanoid prompting agent has on the visual attention, prompt compliance, and task performance of children with ASD during hand-washing activities.

This is the first research of its kind in the field of humanoid robot prompting agent guiding children with ASD through an activity of daily living. Therefore, it is wise to begin with a pilot study, the purpose of which is to show plausibility of the key underlying assumptions of our hypotheses, and to probe what questions are important to be answered later in a more rigorous randomized control trial. For this reason, the pilot study should be exploratory in nature, having a flexible experiment design, and relatively low experiment setup cost.

2.1 Wizard of Oz Experiment Design

The Wizard of Oz (WoZ) is an experiment design widely used in Human Computer Interaction (HCI) and Human Robot Interaction (HRI) researches[REF]. In a typical WoZ study, there is an interactive agent that is not yet fully autonomous, and is remotely controlled by a human wizard (operator), and this fact is concealed from the user being tested until after the study. The wizard may control one or many parts of the agent, such as speech recognition and understanding, affect recognition, dialog management, utterance and gesture generation and so on [REF]. The advantage of a WoZ study is that it does not require a large amount of work spent in implementing the artificial intelligence (AI) behind the agent – it is taken care of by the wizard. This is great for testing hypotheses early on in the design loop, enabling us to obtain feedbacks from users, learn, and iterate through design cycles faster. Of course, care needs to be taken to ensure the mocked up part of the AI is implementable in the near future, since the real purpose of the mock up is to have an early knowledge of the real design constraints, not trying to provide a less constrained solution.

The characteristics of a WoZ study fits our pilot study requirements, where we want to learn early the important design questions regarding building an effective ADL prompting robotic agent for the children with ASD population. Therefore, we will conduct a WoZ study, in which a humanoid robot whose motions and speech are preprogrammed, but the decision and timing of their executions are controlled remotely by the researcher. This is mocking up computer vision algorithms that understands the child with ASD's actions, speech recognition algorithms that recognize the child with ASD's verbal interactions, and the AI decision making algorithms that decides what prompts to deliver and when to deliver them.

During each WoZ study trial, the child with ASD will be asked to complete the hand-washing activity in the washroom with the supervision of one of his/her parents, with the help of the NAO robot, or with the help of both the parent and the robot. The researcher, and the parent if the child is to be assisted only by the robot, will be in an adjacent room out of the child's view to observe his/her hand-washing activity. However, the parent may enter the washroom if the child needs physical assistance to complete a step. A controlling interface running on a laptop, connected wireless to the robot, will be used by the researcher to remotely control the robot, as well as to monitor the progress and responses of the child through the video feeds of the cameras installed in the washroom.

2.2 Recruitment

Participants will be recruited from a previous autism study who indicated that they would be interested in participating in future studies related to the development of the COACH prompting system.

Participants will be children between the ages of 4 to 15 with a diagnosis of ASD, and their parent. Six children will be recruited. This sample size is typical for studies of this nature for children with ASD. For example, a pilot study by Bimbrahw et al. [REF] and a Wizard of Oz study by Bhargava et al. [REF] both involved a similar sample size of the children with ASD in their studies. Another reason we chose six children for this pilot study is to equally explore the two permutations of experimental conditions (i.e. A-B-C and A-C-B, see Section 2.5.2). Participant demographics will be recorded and will include age, sex, and the Social Responsiveness Scale (SRS) test results. The SRS is a commonly used tool to identify the presence and estimate the severity of ASD [REF]. The results of the SRS will allow the research team to substantiate a diagnosis of an ASD for the child participants before proceeding with the study.

The inclusion criteria for enrolling in the study are as follows:

- Boys and girls between the ages of 4-15
- Parent report of a clinical diagnosis of an ASD to be confirmed through administration of the Social Responsiveness Scale (SRS)
- Has difficulty independently completing self-care activities, specifically hand-washing
- Has the ability to follow simple, one-step verbal instructions
- Ethical consent granted by parents or primary guardian
- Does not exhibit severely aggressive behavior

Each participating family will be given a $200 honorarium per child subject upon completion of the study (please see Appendix K Study budget sheet). All participants will be able to withdraw from the study at any time. The honorarium will be adjusted to be proportionate to the number of visits completed (e.g. completing 3 visits means the participated child will receive $100 ($200 * 3 / 6 = $100)). This will be made clear to participants at the time of consent.

2.3 Humanoid Robot NAO

We chose the half-torso version of the commercially available humanoid robot NAO from Aldebaran Robotics as our robotic prompting agent. NAO is a humanoid robot about half a meter high in full torso [FIGURE of full torso NAO]. It is designed by Aldebaran Robotics to primarily serve in educations for children and in academia research in robotics. Because of this, NAO is designed to have a very likable appearance – one with baby like facial features. Also, NAO is equipped with the state of the art mechanical, electrical, embedded, control, and local network communication systems. It also has cameras and sonar sensors for computer vision algorithms for scene understanding, path planning, and obstacle avoidance. The software development kit (SDK) provided is very easy and powerful to program with. Also, an even easier graphic user interface (GUI) for robot behavior programming, the Choreographe software, is also available. One caveat of using NAO for SAR HRI researches is that it is only equipped with a single degree of freedom finger dexterity, though other joints in its body are much more mobile. It is more than enough for doing simple pointing and other non-contact gestural prompts in sync with verbal interactions. It just cannot perform detailed hand gesturing. This makes NAO less capable in demonstrating a hand-washing step in high detail.

From a HRI research perspective, Aldebaran Robotics took care of designing the intrinsics level of HRI, where NAO has a likable appearance and child like neutral gender voice, although it's incapable of facial expressions. The design decisions we face when using NAO for this thesis is on the behavior level of HRI. Design decisions such as voice intonation choice, verbal prompts, motion gestures and gaze, and eye blinks using LEDs in eye regions are made in this thesis. The objective of this thesis is then to ultimately find out if the lower two levels of design decisions made are able to cumulate to the child with ASD perceiving NAO as a role model / supervisor / assistant during hand-washing.

For our pilot study, we use the half-torso version of NAO because we do not require any mobility from NAO – it is fixed on the sink table top [FIGURE]. The relevant functionalities of NAO we will utilize for delivering prompts include:

- Verbal prompting through its bilateral loud speakers on the head and speech synthesis functionality
- Body gesturing through its moving head and arms (although its fingers are not capable of hand gesturing)
- Flashing LEDs on the eyes and ears

2.3.1 Verbal Prompts

We used the text-to-speech engine from NAO to synthesize the verbal prompts. The pitch of NAO's voice is changed to a lower one than default for the verbal prompts to give a more authoritative feeling. The reward verbal prompt remains the default pitch, though, to give an exciting praise. The verbal prompts are worded as short, three or four word phrases, such as "turn on the water" or "rinse your hand", and a pause is put between the action and the subject so that the prompts sound clearer and is easier to understand to children with ASD.

2.3.2 Gesture Prompts

There are several kinds of gesture prompts NAO needs to perform:

- **Attention grabber (AG)**: When prompting is needed but child is not looking at NAO, NAO waves at child to grab the childs attention.
- **Motion demonstrating prompt (MoDemo)**: NAO demonstrates to child the motion of interaction (e.g. turning tap, scrubbing, rinsing, etc.).
- **Object pointing prompt (ObjPt)**: NAO points to the physical object of interaction.
- **Reward (REW)**: After a task is successfully completed, NAO flashes LEDs as a positive reinforcement.

The gaze behaviour of NAO during gesture prompts is also important and is grouped as: looking at child (when delivering AG, MoDemo, REW), and looking at object (AR, ObjPt). The gesture and gaze motions can be programmed using NAO's software, Choregraphe.

2.3.3 Wizard of Oz Remote Control

The WoZ experiment setup involves controlling the robot remotely behind the scene by a human operator, the wizard. A touch screen laptop will be used as the user interface for the operator, and the behaviors of the robot are presented as buttons on the screen, with the camera views displayed along side. Keyboard accelerators are also implemented for faster access to robot actions.

2.4 Surveys

2.5 Protocol and Setup

2.5.1 Entrance Survey and SRS

Prior to their first HomeLab visit, the parent will be asked to complete the Social Responsiveness Scale (SRS) [REFsection]. If the child meets the SRS score (minimum of 76 T-score), the same parent will then be asked to complete the entrance survey before their first visit of the HomeLab. This is to capture the child's demographics, his/her hand-washing ability level and to gather information to help the research team configure the system to the child's preferences [REFsection]. The same parent who has completed the entrance survey should accompany the child through all the HomeLab visits.

2.5.2 Protocol Overview

Each child will visit the HomeLab on the 12th floor of Toronto Rehab Hospital once a week with a total of six visits with his/her parent. The six visits will be evenly divided into three phases. The three phases are the baseline phase (Phase A) and the intervention phases (Phase B and Phase C). In Phase A, the child will be asked to wash hands by him/herself as independently as possible. The parent will be instructed to provide assistance to the child when the parent sees necessary (as outlined below). In Phase B, the child will be assisted by the robot NAO alone. The parent is out of view in the room adjacent to the washroom, and comes into the washroom to prompt when the parent sees necessary. In Phase C, the child will be assisted by NAO and the parent together. The parent remains in the washroom and assists NAO in its prompts.

It will take about an hour to an hour and a half for each visit. The child will be asked to wash his/her hands eight times for every visit, for a total of forty-eight trials per child. This way, the consecutive trials conducted within a day is hopefully not too much for the child to cause fatigue. Also, having sixteen trials per phase gives a sufficient sample size for both quantitative and qualitative (visual) analysis. The child and his/her parent may take short breaks after each hand-washing trials. The break may last as long as the child needs until he/she is willing to continue the trial. If the parent feels the need, they may leave and come back to finish the rest of the day's trials another day. They will not be withdrawn from the study unless requested.

The participating subjects will be randomly assigned one of the two phase orders: A-B-C and A-C-B. This will reduce the confounding effect of learning when we compare between phase B and C.

2.5.3 Experiment Setup

-todo

2.5.4 Specific Protocol

The hand-washing activity will be broken down into seven steps: turn on the water, wet your hands, get some soap, scrub your hands, rinse your hands, turn off the water, and dry your hands. These steps are modified based on Bimbrahw et al. pilot study [REF]. These constitute the same steps as Bimbrahw's except that the first (i.e. turn on the water and wet your hands) and the last step (i.e. turn off the water and dry your hands) are now four individual steps to ensure that each step only involves one action.

Phase A (Baseline Phase) The first phase will be the baseline phase and will include sixteen trials of hand washing. The child will be asked to complete the hand-washing as independently as possible. During this phase, the parent will be present in the washroom while the child is completing the hand-washing steps. The parent will verbally and/or physically assist and give positive reinforcements to the child whenever the parent feels necessary.

Phases B and C (Intervention Phases) The rest of the two phases will be the intervention phases and will include sixteen trials of hand washing each. The child will be asked to wash his/her hands with the help of NAO or of both NAO and the parent in the washroom. During each trial, NAO and the parent will wait for the child to start each step. If the child has trouble, an appropriate prompt will be delivered from NAO in order to help the child complete the step. If the child does not respond to NAO's prompt, an attention grabber will be delivered to capture the child's attention from the prompting agent. The attention grabber may be repeated for the second time to the child if he/she fails to respond to it. A verbal reward will be delivered to the child once he/she completes the step.

The parent's role in phase B and C differ in that, in phase C, the parent takes a more active role to prompt the child of what to do by standing next to the child. On the other hand, in phase B, the parent takes more of a back seat role, being out of view and coming in to prompt only for the purpose of reminding the child to listen to the robot, but leaves the specific step to be prompted by the robot. Of course, if the child doesn't respond to any of the prompts, the parent will need to physically intervene and complete the step together, just like in phase A. After the physical intervention, the parent will then instruct and encourage the child to continue the rest of the hand-washing steps on his/her own by following the robot.

There are three prompt categories that the NAO robot will deliver when interacting with the child (please see [TABLE] for the specifics of each prompt used):

1. **Step Prompt** (to prompt the child through a hand-washing step):

   A verbal prompt will be delivered, such as Please [step name] (e.g. Please turn on the water.). Synchronous to the verbal, a visual prompt will also be delivered. This is a two-part gesture prompt of: first, demonstrating the motion of interaction while looking at the child (MoDemo); second, pointing to the sink object (e.g. the tap) while looking at the object (ObjPt). A maximum of two prompts will be given to the child. If the child does not respond to the second prompt or has started the step but does not complete the step within the a reasonable time, the parent will be asked to help the child complete the step.

2. **Attention Grabber** (to catch the child's attention to the NAO robot or the avatar):

   A verbal prompt will be delivered, such as Hi, [child's name]! Synchronous to the verbal, a visual prompt will also be delivered. This is an attention grabbing gesture of waving and looking at the child (AG). A maximum of two attention grabbers will be given to the child in order to get his/her attention to look at the robot/avatar. The parent will be asked to instruct the child to look at the robot/avatar if he/she does not respond to the second attention grabber.

3. **Reward** (to provide positive reinforcement when the child attempts a step without the help from his/her parent):

   A verbal reward (i.e. Great!) will be delivered while looking at the child and switching back and forth the colors of the light-emitting diodes (LEDs) on the eyes after successfully performing a step (REW).

For each trial, in addition to the three prompt categories stated above, the NAO robot will also deliver a short introduction before the start of each trial, a re-intro after the parent finished assisting the child through a step, and an outro at the end of each trial. The introduction is a two-part prompt. The first part is an attention grabber. The second part consists of a verbal prompt (i.e. Let's start washing hands.) with a simple conversational gesture. The re-intro is a verbal prompt (i.e. Let's continue washing hands.) with a simple conversational gesture. Same as the introduction, the outro is a two-part prompt. The first part consists of a verbal prompt (i.e. Good job, [child's name]!) with a gesture of fist pumping in the air. The second part consists of a verbal prompt (i.e. You are all done.) with a gesture signifying all the hand-washing steps have been done.

2.5.5 Post-intervention Survey and Exit Survey

During the last visit, the same parent who has completed the entrance survey will be asked to fill out the post-intervention survey and the exit survey [REFsection], which will allow him/her to provide the research team with his/her feedback regarding the device. A variation of the post-intervention survey will be verbally administered by the researcher to the child participant to capture his or her views of the system [REFsection]. This information will be used by the research team to better understand which aspects of the system are effective, which are not, and how, if in any way, the system should be changed.

2.5.6 Data Collection

All phases will be video recorded by the overhead, the scene, and the Kinect cameras and will be audio recorded by the microphone from the scene camera. The overhead and scene video data will be reviewed and annotated by two annotators. The inter-rater reliability will be calculated using Cohen's Kappa [14]. The overhead video data will be used to score the participants' prompt compliance and hand-washing performance. The scene video data will be used to evaluate the participants' engagement during the whole activity. The effect of embodiment on engagement, compliance, and performance will then be explored qualitatively and quantitatively.

The Kinect video data will not be annotated. Instead, it will be used to evaluate the automatic gaze estimation algorithm that we developed. Specifically, the Kinect video data will be used by the gaze estimation algorithm as input and the output predictions will be compared with annotations of the scene video data to derive the algorithm's prediction accuracy.

2.5.7 Ethics

The WoZ is approved by the Research Ethics Board (REB) of University Health Network (UHN), belonging to which is the Toronto Rehab Hospital, where the study is conducted.

Consent and Assent Participants will be given a package of consent/assent forms prior to starting the study [REFsection]. One of the parents will need to provide their consent for their child and themselves to participate in the study. In addition, child participants will need to provide their assent to participate in their every visit of the study.

Interested families will receive an information/consent package (please see Appendices F to H) prior to starting of the study. This package includes consent/assent forms for participation in the study for the parent and child with ASD (these forms include study details and research contact information) as well as consent to be videotaped for the parent and child with ASD. Consent from the parent and assent from the child with ASD will be given if and when they feel comfortable that they understand the information presented. Potential participants of both parents and children will have up to a week to decide if they would like to participate, although they may consent to participate as soon as they feel comfortable doing so. Parents will need to provide their consent for their children (please see Appendix F) and themselves [REFsection] to participate in the study. In addition, child participants will need to provide their assent in their every visit of the HomeLab during the study (please see Appendix H) to participate. Parents will be required to consent to having their children and themselves videotaped during the study. The parents will be informed that they and their children may withdraw from the study at any time without penalty.

Confidentiality Each participating family (parent and child with ASD pair) will be assigned a code number when they sign the consent/assent. All data in the study will be labelled with these code numbers only - the names of the participants will appear only on the information and consent/assent forms and will be kept confidential. Consent forms will be placed in a secure and locked area in the PI's laboratory, with access exclusively restricted to the research team. All forms will be destroyed seven years after the study publication.

The information and data collected will remain strictly confidential and will not affect any of the participants (both the parent and the child)' employment, care, or treatment in any way. A code number will be assigned to each parent and child participant when they give consent. This code number, instead of their name, will be used for all data collection and analysis. Direct quotes may be included in the final research paper but names will not be used in any report or publication. Privacy of participants (both the parents and the children) will be ensured by omitting all participant information from participant data, by employing data encryption, and by storing data on a secure server. If and only if participants consent, participants (both the parents and the children) video data may be presented for educational purposes. If any images or videos are used in presentations and publications, faces and other identifiable features will be masked.

Both the video and audio data will be stored temporarily on the touchscreen laptop's hard drive during each child's visit. The data will be encrypted and transferred to the TRI servers as soon as after each child's visit. The portable devices, such as USB sticks, will be used to transfer the data to the TRI servers. All files stored in the portable devices will be password protected and encrypted. The data on the laptop's hard drive and the portable devices will then be purged immediately after transfer.

Data Storage All soft (electronic) data will be encrypted before any transfer is made. All data will be password protected and be stored on the TRI servers with access restricted to the research team. The laptop used for the study will be password protected so that only the research team has the access to it. All computerized data will be password protected. All survey data will be stored in a locked cabinet different from where the consent forms are stored. Access to all the data will be restricted only to the supervisor and researchers involved in the project.

After the study is completed and the results of the study are published, data will be stored for at least seven years from study closure. All data will be destroyed seven years after the study closure. Data contained on paper material will be destroyed by shredding the material. Data contained on electronic media will be destroyed by erasing or other removing the data in such a way that it cannot be retrieved.

2.6 Measures

2.6.1 Video Data Measures

To evaluate the effectiveness of the robot prompts on child's step completion, to measure the child's compliance to the prompts, and to investigate their relationships with the child's engagement level during hand-washing, the following metrics are calculated and analyzed from the video data annotations for each trial:

Prompt Effectiveness

- **Total Number of Incomplete Steps**: the number of hand-washing steps that were prompted but the child failed to attempt or attempted but failed to complete.
- **Total Number of Parent Prompts**: the number of prompts delivered by the parent (e.g. verbal, pointing, motion demonstrations, nudging, guiding, and physically intervening).

Responses to Prompts

- **Compliance Rate**: the percentage of prompts that the child followed correctly.

- **Not Affected By Prompt Rate**: the percentage of prompts that the child ignored.

Engagement and Visual Attention

- **Total Number of Times Child Smiles**: the number of prompt sections that the child smiled.

- **Total Number of Times Child Murmurs**: the number of prompt sections that the child murmured.

- **Looking at Prompting Agent Rate**: the number of prompt sections that the child looked at parent when parent was prompting or at robot when robot was prompting.

The specific definitions for each measure will be discussed further in the analysis section 2.8.3. The difference between prompts and prompt sections is discussed in 2.7.1.

2.7 Video Annotations

In order to calculate the final measures outlined above, intermediate measures need to be extracted from the video data. This is the process of video annotation.

2.7.1 Annotation Framework

Only the scene camera videos will be annotated, since this view alone suffices in informing both the progress of the child in hand-washing steps and the child's response to prompts.

Each video file usually contains one hand-washing trial, sometimes two. The annotator needs to scroll through each video until the scene of the child entering the washroom, marking it as the start of a trial. The child leaving the washroom marks the end of a trial.

A trial contains many hand-washing steps, and for each step, the parent and/or the robot may give several prompts. For consistency and convenience, the annotator divides the video into segments we call "prompt sections", and describe each prompt section using a 3-part scheme. The first part describes the child's actions before any prompts, the second describes the prompting agent's prompts, and the third describes the child's actions after the prompts. The intermediate measures to be annotated in each part of the prompt section are shown in Table 2.2.

Intermediate Measure	**Type**	**Description**
Step	Nominal	Prompting step, 0 no step, 1 intro, 2 turn on water, 3 get soap, 4 scrub hands, 5 rinse hands, 6 turn off water, 7 dry hands, 8 all done, 9 wet hands
Child's Action Before Prompts		
Time Start	Ordinal	Time stamp for start of prompting section
Time Stop	Ordinal	Time stamp for end of prompting section
Attempted Step Before Prompt	Nominal	
Attempted Step Successfully Executed Before Prompt	Ordinal	0 incomplete, 1 complete but low quality, 2 complete with high quality
Prompting Agent's Prompts		
P Verbal	Ordinal	Parent verbal prompt, 0 no verbal prompts, 1 prompt for compliance to robot, 2 prompt for step
P Gesture	Ordinal	Parent gestural prompt, 0 no gesture prompts, 1 quick point, 2 sustained point, 3 motion demonstration, 4 motion demonstration and point, 5 nudge, 6 guide arm, 7 do step fully
P Reward	Boolean	Does parent give reward
R Verbal	Ordinal	Robot verbal prompt, same coding as P Verbal
R Gesture	Ordinal	Robot gestural prompt, same coding as P Gesture
R Attention Grabber	Boolean	Does robot give attention grabber
R Reward	Boolean	Does robot give reward
Child's Action after Prompts		
C Looks At P/R	Nominal	Child looks at the prompting agent, 0 no looks, 1 looks at parent, 2 looks at robot, 3 looks at both
C Smiles	Boolean	Does child smile
C Murmurs	Boolean	Does child make a verbal sound
Attempted Step After Prompt	Nominal	
Attempted Step Successfully Executed After Prompt	Ordinal	Same coding as Attempted Step Successfully Executed Before Prompt
Attempted Step Is Correct Although Different From Prompt	Boolean	Is this one of those times that the prompts are wrong or ambiguous and child's actions make sense despite different
Number of Prompts Till C Executes Correct Step - Parent	Cardinal	Count number of prompts as any distinct actions performed by the parent before child executes the correct step.
Number of Prompts Till C Executes Correct Step - Robot	Cardinal	Same as above, but counting robot prompts

Table 2.1: The intermediate measures annotated from the videa data

A hand-washing step could have multiple prompt sections. Take for example the following scenario: The child executes the wrong step before prompts, so the parent prompts the correct step, but the child ignores the prompt and continues the wrong step. This constitutes one prompt section. Then the parent prompts again, and the child finally follows the prompt and executes the correct step. This constitutes then another prompt section. For this example, because the parent prompts a second time without waiting for the child to stop his/her current action, the second prompt section should have a blank for the Child's Action Before Prompts.

A hand-washing step could also have multiple prompt sections because of the step's nature. For the "extended steps" (i.e. scrubbing, wetting, rinsing, and drying), even when the child is executing the correct step, the prompting agent may deliver more prompts to encourage the child to keep doing the same step for an extended period of time. This is in contrast to the non-extended steps (e.g. turning on the water), where a single action from the child marks the completion of that step. An example of an extended step with multiple prompt sections is: The child starts rinsing before prompt, then the parent tells the child to keep rinsing. The child continues to rinse. The parent says again "keep rinsing", and the child rinses more and then decides to stop. This constitutes one prompt section. Then, the parent prompts to rinse more again. The child follows. After a while, the parent decides this is enough and prompts for the next step. This marks the end of the second prompt section. For the first prompt section, it contains two prompts from the parent. This is intentional, for the purpose of convenience – we group any consecutive prompts (can be from either the parent, the robot, or from both) resulting in the same actions from the child as one prompt section. This grouping does not affect any of our measures for Prompt Effectiveness or Responses to Prompts, since those measures count the number of steps or prompts, not prompt sections. However, this grouping does affect the measures in Engagement and Visual Attention, since these measures count the number of prompt sections instead.

2.7.2 Annotation Tools

The videos are played back by the software Media Player Classic - Home Cinema (MPC-HC 64-bit v1.7.8), where timestamps of millisecond resolution can be obtained. The annotations are recorded onto Microsoft Office Excel spreadsheets, and each sheet exported to CSV files to be analyzed.

2.7.3 Annotators and Inter-rater Agreement

- number of annotators - percentage of overlap - inter-rater agreement calculation (method, what's good enough)

2.8 Data Analysis and Results

2.8.1 Participants Recruited

Due to limitation of time, we were only able to recruit one subject. Our participant is a thirteen years old male child of Asian ethnicity. He was accompanied to the trials by his mother, who was the one that answered all surveys.

Child's Demographics and Inclusion Criteria Fit

The child has been clinically diagnosed of Autism Spectrum Disorder. We also conducted the Social Responsiveness Scale Survey, and he obtained a T-score of 79, passing the minimum score for severe ASD. Through the Entrance Survey, we learned that the child has difficulty independently completing self-care activities (a 2 on a scale of 1 (not independent at all) to 5 (completely independent)), and this includes hand-washing (also a 2 on the same scale). We also learned that the child is able to do but not good at verbal communication (a 3 on a scale of 1 (very not well) to 5 (very well)). Specifically, the child can only speak one or two words at a time to express what he wants, uses iPad for communication, and often just murmurs illegibly. However, the child has the ability to follow simple, one-step verbal instructions (a 4 on a scale of 1 (very not well) to 5 (very well)). Lastly, the child does not exhibit severely aggressive behavior (a 1 from a scale of 1 (never) to 5 (often)). The above shows that the child fits in our inclusion criteria.

2.8.2 Experiment Design Change

We were not able to counterbalance the confounding effect of learning through randomly assigning participants to phase orders (A-B-C versus A-C-B), since we only recruited one participant. Instead, we decided to control it by splitting the phase B into two segments, one before phase C and one after. Thus, we conducted the study in the following phase order: A-B-C-B (i.e. parent alone phase - robot alone phase - robot parent phase - 2nd robot alone phase). This way, we can compare phase B and second phase B to see how much does learning in phase C affect our results. Also, phase B and second phase B won't have sixteen trials each due to limitation of time. Instead, we conducted these two phases only long enough to see a stable response. As a result, we had 16 trials for phase A, 8 trials for phase B, 21 trials for phase C, and 5 trials for second phase B. Note that the intervention conditions of phase B and second phase B are meant to be the same.

2.8.3 Video Data Analysis and Results

Analysis Method

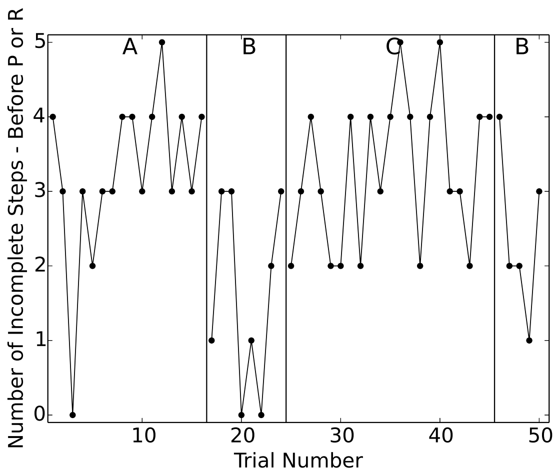
Prompt Effectiveness

To reflect how effective our prompting system is, we show whether the system can reduce both the number of incomplete steps and the number of parent prompts.

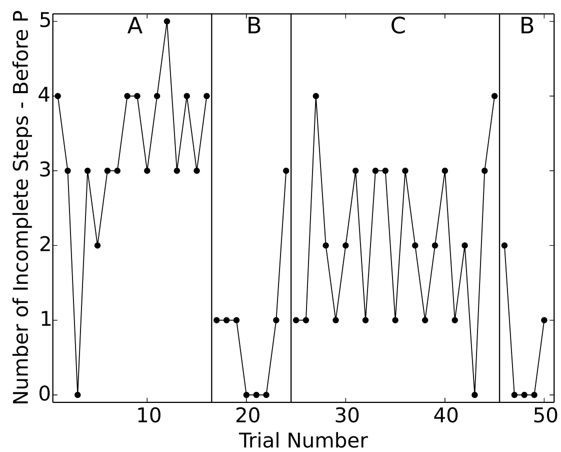
Number of Incomplete Steps We assumed that prompts can be ordered by their level of authority over the child, with robot prompts the lowest level of authority, followed by parent non-physical prompts (e.g. verbal prompts and gestures such as pointing and motion demonstrations), and lastly parent physical prompts (e.g. nudging, guiding the arm, and completely executing the step for the child). Figure 2.1 shows a series of plots for the measure "Total Number of Incomplete Steps", differing in the prompt levels threshold used to produce the figure. The prompt levels threshold defines what steps were counted as incomplete when plotting the figures, e.g. Plot 2.1a ("before parent and robot prompts") counts any steps that were prompted by either robot or parent as incomplete. These four plots show a progression of allowing more and more steps to count as complete by removing lower prompt levels from the threshold. For example, the next Plot 2.1b ("before parent prompts"), allows steps prompted by the robot to also count towards completed steps, and Plot 2.1c (”before parent physical prompts”) allows steps prompted by parent’s non-physical prompts to also count towards completed steps, and lastly, Plot 2.1d (”overall”) counts every completed steps even if they were prompted by parent’s physical prompts.

By comparing the plots from one to the next, the effects of each newly added prompt level is apparent. The most important comparison is between Plot 2.1a (”before parent and robot prompts”) and Plot 2.1b (”before parent prompts”), demonstrating the effectiveness of the robot’s presence. We see that parent alone phase (phase A) was unaffected since robot wasn’t present, but introducing the robot in the rest of the phases show effectiveness: robot alone phase (phase B) from 1.5 to 0.5, robot parent phase (phase C) from 3 to 2, and robot alone repeat phase (second phase B) from 2.5 to 0.5.

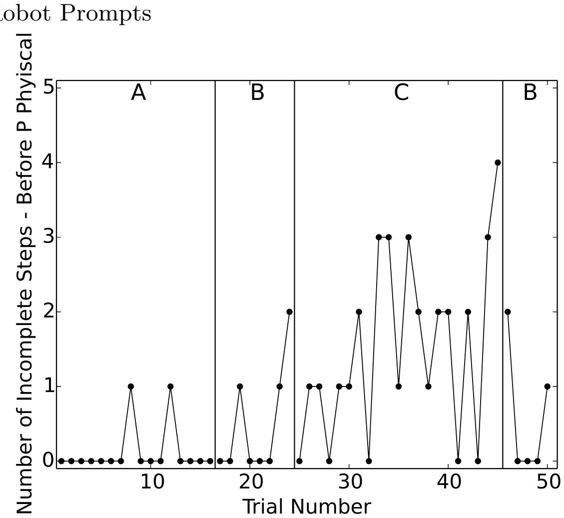
Comparing Plot 2.1b (”before parent prompts”) against Plot 2.1c, we see the effectiveness of parent non-physical prompts: parent alone phase moved from 3 to 0, robot alone phase from 0.5 to 0, robot parent phase from 2 to 1.5, robot alone repeat phase is unaffected.



(a) Total Number of Incomplete Steps - Before Parent or Robot Prompts

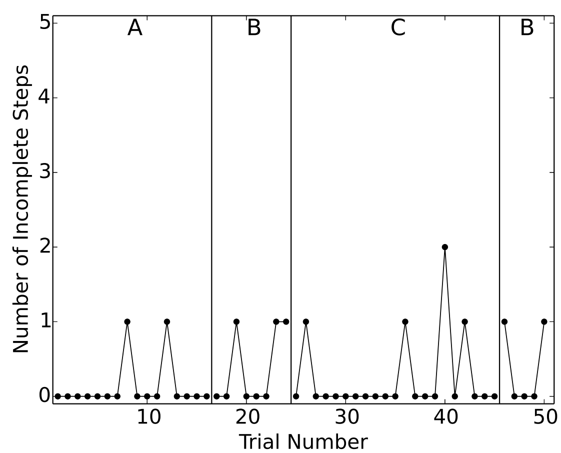


(b) Total Number of Incomplete Steps - Before Parent Prompts



u

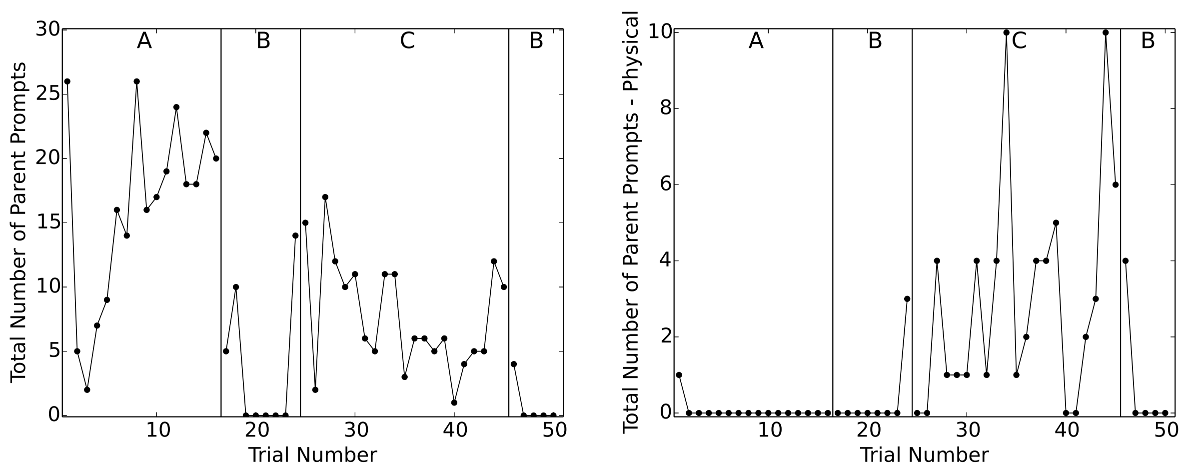
(c) Total Number of Incomplete Steps - Before Parent Phyiscal Prompts



(d) Total Number of Incomplete Steps - Overall

Figure 2.1: Total Number of Incomplete Steps

Number of Parent Prompts The measure "Total Number of Parent Prompts" is plotted in Figure 2.2. Plot 2.2a is for the overall count (i.e. counting both physical and non-physical prompts). It shows that during parent alone phase (A), the measure has an upward trend from 5 moving to 20. However, in robot alone phase (B), we have a sudden drop leveling at near zero. In robot parent phase (C), the measure has a downward trend moving from 15 to 5. In robot alone repeat phase (second phase B), we again observe a near zero level. By comparing the measures across phases, we see that the robot's presence were effective in reducing the number of parent prompts, especially in robot alone and repeat phases. Plot 2.2b is for the physical prompt count. This plot shows when the parent resorts to a higher prompt level (i.e. physical prompts such as nudging, guiding, and physically intervene) in order to get the child's compliance. We see that the level is mainly near zero for all phases except for robot parent phase (C), leveling around 2.5.



(a) Total Number of Parent Prompts - Overall (b) Total Number of Parent Prompts - Physical

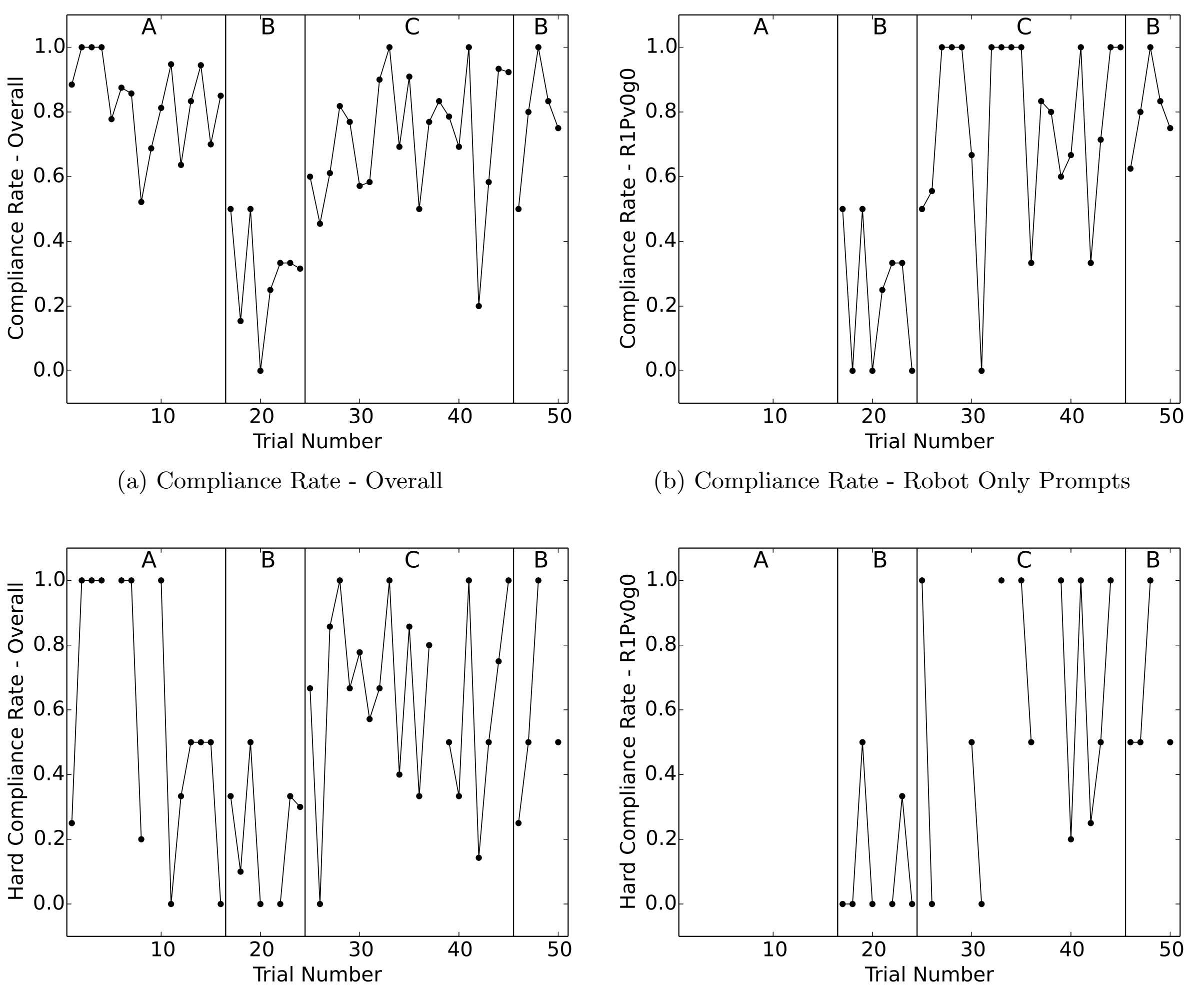
Figure 2.2: Total Number of Parent Prompts

Child's Response to Prompts

To illustrate the child's different responses to the prompts, we characterized child's responses into three categories: "compliance", "not affected by prompt', and others.

Compliance Rate A response is counted towards "compliance" if the child executes the correct step in response to the prompt. If the child was executing the wrong step before prompt, and is converted into doing the correct step due to prompt, we call this hard compliance. The compliance and hard compliance response rates are shown in Figure 2.3.

Plot 2.3a shows the overall compliance rate, with parent alone phase (A) leveling at 80%, robot alone phase (B) leveling at 30%, parent robot phase (C) moving upward from 60% to 80%, and robot alone repeat phase (second phase B) leveling at 80%. We see that when the robot was first introduced in robot alone phase (B), the child did not comply to the prompts. However, by going through phase C where the parent prompts for child to follow the robot, the child complies more readily in the robot alone repeat phase, achieving similar level of compliance as the parent alone phase. We need to note that this plot includes prompts delivered by the robot, by the parent, and by them together. Even in robot alone and repeat phases, the parent still comes into the washroom and prompts when the child isn't complying to the robot. To see whether the robot alone can potentially guide the child through the whole hand-washing activity with minimal parent involvement, we plotted the compliance rate counted over only prompts delivered by the robot, shown in Plot 2.3b. This plot confirms the levels observed in the overall plot, validating the improvement of compliance rate seen in R Alone Rep phase. To investigate to what extent the child is compliant, the overall hard compliance rate is shown in Plot 2.3c, with parent alone phase (A) split leveling at 100% and 35%, robot alone phase (B) leveling at 25%. The robot parent phase (C) averaging around 60% but the spread increases as trials went on. Lastly, the robot alone repeat phase (second phase B) levels at 50%. Similar to above, we observe an improvement of hard compliance between robot alone and repeat phases. Looking at the robot prompts only Plot 2.3d, the robot alone phase (B) drops to almost 0%, while robot alone repeat phase (second phase B) remains at 50%.



(a) Compliance Rate - Overall

(b) Compliance Rate - Robot Only Prompts

(c) Hard Compliance Rate - Overall

(d) Hard Compliance Rate - Robot Only Prompts

Figure 2.3: Compliance Rate

Not Affected By Prompt Rate A response is counted towards "not affected by prompt" if the child was executing a wrong step and did not change after the prompt or was idling and did not change after the prompt. The not affected by prompt rate is shown in Plot 2.4. We see that for most phases, it levels at 15%, but for robot alone repeat phase (second phase B) it is at 35%. This shows the robot prompts were ignored more when the robot was first introduced, but improves to an acceptable level through phase C.

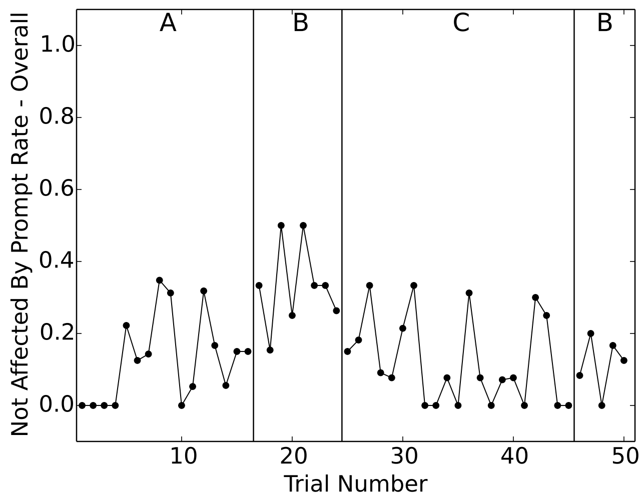


Figure 2.4: Not Affected By Prompt Rate

Engagement and Visual Attention

To further characterize child's response to different prompting phases, we investigate how many times the child smiles and murmurs during step execution, and how often the child looks at the prompting agent during prompting and step execution.

Number of Times Child Smiles The measure "Total Number of Times Child Smiles" is shown in Plot 2.5. In it, parent alone phase (A) levels at 1.5, robot alone phase (B) levels at 0.5, robot parent phase (C) has a large spread and averages around 3, and robot alone repeat phase also has a large spread and averages around 4. It shows that the child smiles much more in later phases compared to earlier phases, and particularly, smiles in the repeat phase more than in robot alone phase.

Number of Times Child Murmurs The measure "Total Number of Times Child Murmurs" is shown in Plot 2.6. In it, parent alone phase (A) has a large spread, averaging around 4. Robot alone phase (B) levels at 0.5. Robot parent phase (C) has a large spread, averaging around 4. Robot alone repeat phase (second phase B) levels at 2. It shows that the child murmurs much more often when the parent is present. Also, child murmurs in the repeat phase more than the robot alone phase.

Looking at Prompting Agent Rate A prompt can be given by the parent, by the robot, or by them together. During prompting and during step execution, the child may turn and look at the parent and/or the robot. The gaze behavior of the child is shown in Figure 2.7 for all the cases above. Because

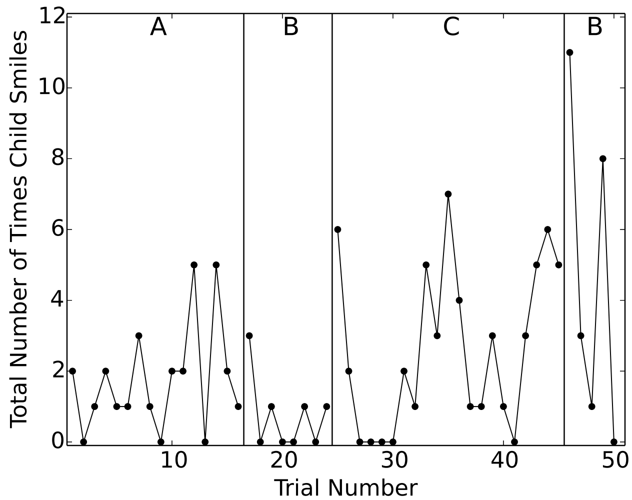


Figure 2.5: Total Number of Times Child Smiles

not all cases have the same amount of data in all phases, we will only mention here the phases that have enough evidence. In Plot 2.7a, ”Looking at Parent Rate - Given Parent Prompted” levels at 40% in parent alone phase (A). In Plot 2.3d, ”Looking at Robot Rate - Given Robot Prompted” trends downward from 50% to 30% for robot alone phase (B), levels at 30% with high spread for robot parent phase (C), and levels at 20% for robot alone repeat phase (second phase B). We see that when the parent and the robot prompt individually, the parent has a higher chance of getting the child's visual attention. Although the robot had similar levels of attention when was first introduced, it dropped as the study went on. In Plot 2.7c, ”Looking at Parent Rate - Given Both Prompted” averages around 50% with high spread in robot parent phase (C). In Plot 2.7d, ”Looking at Robot Rate - Given Both Prompted’ levels at 15% for robot parent phase (C). We see that when the parent and the robot prompt at the same time, the parent had a greater amount of visual attention. It is interesting to note that the parent had similar levels of visual attention when prompting alone and when prompting with the robot. The robot, however, experienced a decrease in visual attention level when the parent prompts with it.

2.8.4 Annotation Inter-rater Agreement Analysis and Results

- method - result

2.8.5 Survey Data Results

Entrance Survey

Child's Experience with Technologies The child is more of a visual learner. He uses a computer at home, and likes to use it very much. He also likes to use other technologies (e.g. iPhone, iPad), and likes to watch movies and TV. He doesn't have a robot toy to play with at home or at school, so the parent doesn't know how much he likes to play with robot toys. The parent never used technologies to help the child with self-care activities except using pictures to teach step by step hand-washing.

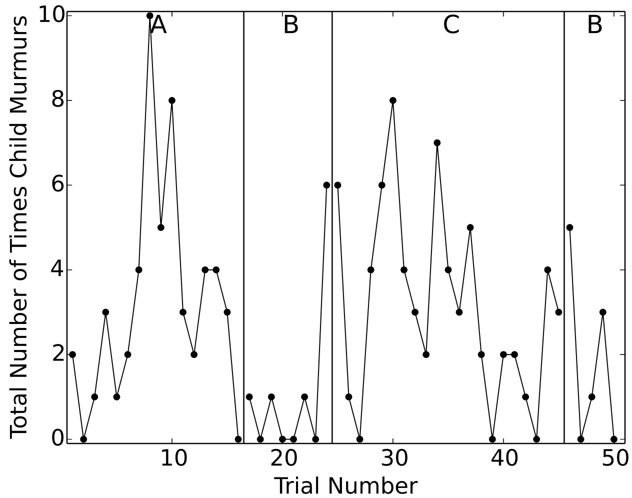


Figure 2.6: Total Number of Times Child Murmurs

Child's Personal Preferences The child is sensitive to sound. He likes Disney cartoon musics, and likes to watch his favorite cartoon scenes repeatedly on the iPad. To reward the child after a good behavior, the parent suggested the following rewards: give extra time to play on iPad, give him praises (e.g. good job), give children books to read and animal dinosaurs to play with, give the parent's iPhone since his favorite musics are on there.

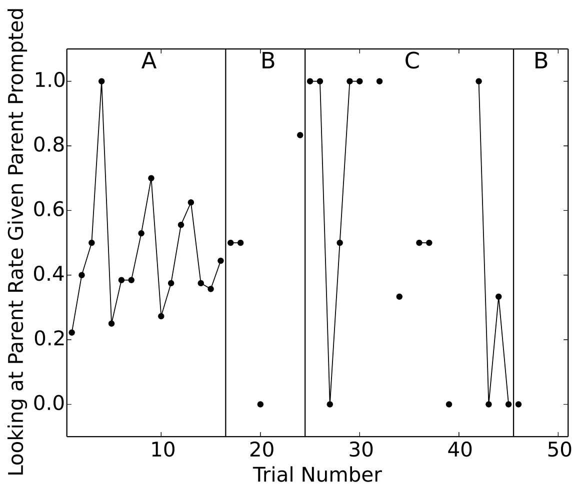
Child's Abilities on Hand-washing and on Other ADLs The parent agrees that the child usually gets distracted when performing hand-washing. To assist him, the parent mainly reminds him to put soap, rinse properly, and dry properly with towel. He needs more prompting in these areas since he always washes in a hurry.

Other activities the child needs help with include: tooth brushing – 2 hours/week; bathing – 4.5 hours/week; dressing – usually just hand the clothes to him, he knows how to put them on, but needs reminders of the order of the clothing, 7 hours/week.

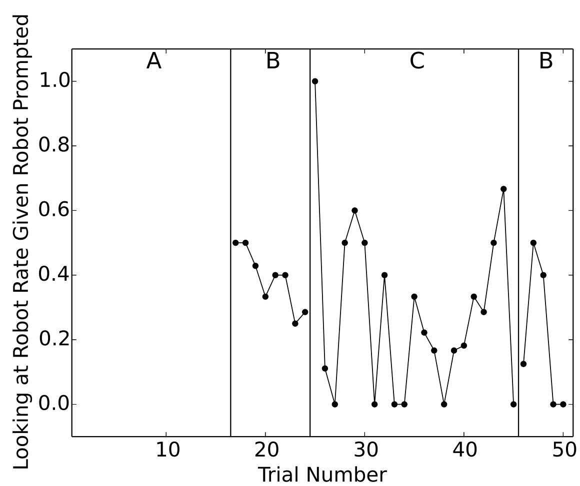
Parent's Expectation and Concerns The parent expected the robot to be helpful in reminding the child to put soap, rinse and dry more, similar to the role of the parent. Some concerns the parent has include: the child may be wondering why does he need to wash hands so many times repeatedly; the child performs well in his comfort zone with the same environment, so it takes a while for the child to get used to the lab environment.

Post-Intervention Survey for Parent

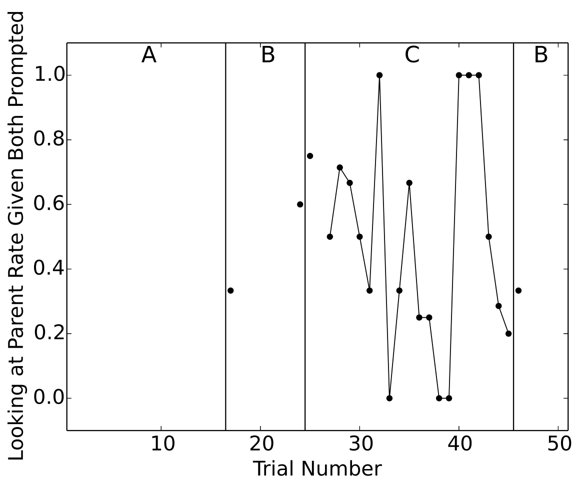
- how do we report the survey?!



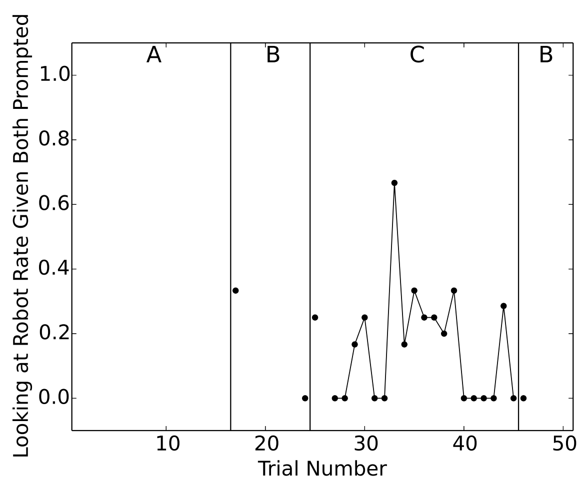
(a) Looking at Parent Rate - Given Parent Prompted



(b) Looking at Robot Rate - Given Robot Prompted



(c) Looking at Parent Rate - Given Both Prompted



(d) Looking at Robot Rate - Given Both Prompted

Figure 2.7: Looking at Prompting Agent Rate

Question No.	**Description**	**Paernt's Answer**
1	Hand-washing steps break down was appropriate	Strongly Agree
2	My child understood the verbal prompts	Agree
3	Robot's verbal prompts were appropriate	Strongly Agree
4	The prompt wordings were similar to mine	Strongly Agree
5	The prompt voice and tone were appropriate	Strongly Agree

Table 2.2: The intermediate measures annotated from the videa data

2.9 Discussion

A Wizard of Oz study has been conducted following the A-B-C-B design. In phase A, the parent alone prompted the child through hand-washing steps. In phase B, the robot alone prompted the child. In phase C, the parent and the robot jointly prompted. Visual analyses were used to analyze the measures from annotated video data of the trials.

We saw that having the robot present in the washroom contributed positively in reducing the Number of Incomplete Steps. Comparing with the parent's non-physical prompts' contribution, the robot prompts had a greater contribution than the parent in all phases that included the robot (both B phases and C phase) (note that we only compare robot prompts against parent non-physical prompts because the robot is incapable of delivering physical prompts). Further more, in the second phase B, the contribution from robot is almost (but not quite) as big as parent non-physical prompts' contribution in phase A (a reduction of 2 instead of 3 incomplete steps). It is interesting that this happened in second phase B, but not in the first. We believe this is mainly due to the child learning to follow the robot after going through phase C, where parent jointly prompted with robot, telling the child to follow the robot's prompts. Although robot prompts in second phase B did not arrive at the same effectiveness as that of the parent's in phase A, we believe that, given sufficient phase C training, the robot potentially could replace the parent as an independent prompting agent.

Another way to look at the robot prompts' effectiveness is this: We first observe that the Number of Incomplete Steps - Overall was always kept to near zero. This is done through either parent prompts or robot prompts or joint. In addition, we saw that the Number of Parent Prompts - Overall was sharply reduced in both of B phases (when robot prompted alone) and experienced a continual reduction in phase C when robot and parent jointly prompted. This shows two things: First, in the two B phases, the robot was able to guide the child through hand-washing without constant parent supervision. Second, in phase C, when the parent jointly prompts with the robot to have the child better follow the robot, the parent was able to gradually fade out the prompts without increasing Number of Incomplete Steps since the child listened to the robot more and more.

One major limitation of evaluating robot prompt effectiveness based on the above measures (i.e. Number of Incomplete Steps and Number of Parent Prompts) is that we could have a trial where the child may be ignoring robot prompts and skipping steps, and yet the trial may still have low Number of Incomplete Steps and low Number of Parent Prompts. This is mainly due to the following limitations: The measure Number of Incomplete Steps is counted only when steps were prompted and were not completed. So if the child largely ignores the prompts or doesn't comply, and skips steps, some of the skipped steps were never prompted, and thus never counted as incomplete. In that case, we might still have a low Number of Incomplete Steps. To worsen the problem, our study did not employ a strict hand-washing steps ordering that the child must adhere to. Also, we left it up to the parent's discretion to intervene as they think the child needs. As a result, the measure Number of Parent Prompts does not fully reflect child skipping steps either. One way to circumvent this problem may be to redefine the measure Number of Incomplete Steps as number of steps that were skipped or failed, regardless of whether the steps were prompted. It seems that child with ASD, or at least the one participant we had, has strict preference on the order of hand-washing steps, and the order is not known a priori and may be different across individuals. Thus, having a flexible prompting scheme in terms of step orders is preferred. However, this means that the decision of what steps need to be executed, and hence which steps were missed, often depends on the sequence of steps executed so far. This makes it troublesome to define what is meant by missed steps, possibly making the analysis process unfeasible. Another way to circumvent this problem is to use other measures that reflect the child skipping steps, or not following the step order prompted – the Compliance Rate and Not Affected By Prompt Rate measures can be used for this purpose.

We saw that both Compliance Rate and Hard Compliance Rate are high in Phase A, C, and second Phase B, and low in first Phase B. And the contrary pattern exists for Not Affected By Prompt Rate. This means that the prompts in first Phase B were complied much less often, and a large portion of that noncompliance was exhibited as the child ignoring the prompts and kept doing some other step or not doing any step at all. At the same time, both Number of Incomplete Steps - Before Parent or Robot Prompt and Number of Parent Prompts - Overall are low in first Phase B. This is precisely the scenario talked about earlier, where the child skips many steps. Thus, it is apparent that the robot was only effective in second Phase B, but not in the first phase B. We believe the difference between first and second Phase B is that second Phase B is after phase C, and first Phase B is before phase C. Phase C functioned as a training phase and helped the child comply better to robot prompts.

There are also other differences that existed between first and second Phase B, so training phase C may not be the reason (or may not be the sole reason) for improved effectiveness of robot prompts in second Phase B. One important change happened to the way the robot was controlled starting at the seventh trial of the first Phase B. We found out, after introducing the robot to the child and conducted six trials, that the child did not follow the order of steps prompted by the robot. This meant the robot needed to change the order of steps being prompted on the fly quickly. However, due to the way its remote control was programmed, the researcher cannot change prompts quickly enough, resulting in the child growing impatient and ignored the robot prompts altogether. Thus, the robot remote control scheme was changed from being inflexible in step order into one that the researcher can select the current prompting step on the fly. Also, instead of implementing a timer, and automatically issuing the prompt after a predefined seconds of waiting, the timing of prompt delivery was changed to be decided by the researcher on the fly. This makes the robot behavior much more flexible and relevant to the child's behaviors. This robot control scheme change affected the last two trials in the first Phase B, and all trials in the second Phase B and Phase C. This change, however, did not cause any abrupt change in any of our measures (comparing the last two trials of first Phase B with the previous trials in the phase). Thus, we believe the robot control scheme change was not the cause of the improved effectiveness of robot prompts in second Phase B.

It is important to mention here that the parent did not prompt the same way uniformly through Phase C, but instead was involved in various degrees. The level of involvement was left up to the parent's discretion for each trial, with the ultimate goal to fade the involvement to minimal by the end of Phase C. The parent involvement in Phase C can be categorized into five levels:

- **Level 0**: the parent is outside of the room out of the child's view, only comes into the washroom to prompt when the child is not following the robot (this behavior is similar to that during phase B)
- **Level 1**: the parent stands beside the child in the washroom, and mainly uses gestures to prompt the child to follow the robot and to demonstrate the step's motions after robot prompts
- **Level 2**: the parent stands beside the child in the washroom, and uses both verbals and gestures to prompt the child about each hand-washing steps, in competition to the robot prompts
- **Level 3**: the parent stands behind the child, and physically guide the child to follow the robot prompts

- **Level 4**: the parent stands beside the child, and prompts as the parent sees fit each hand-washing steps, with no robot present

The Parent Involvement Level is plotted in Figure 2.8. We see that, in Phase C, the parent mainly alternated between trial segments of level 1 and level 3 to train the child to follow the robot, while mixing single trials of level 0 to monitor the child's progress when left with robot alone.

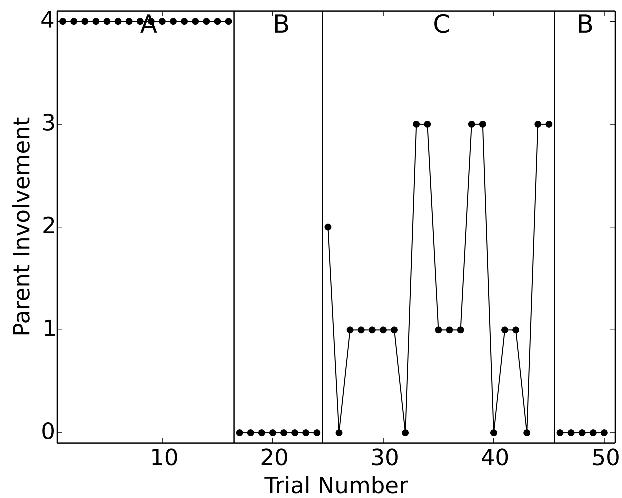


Figure 2.8: Parent Involvement

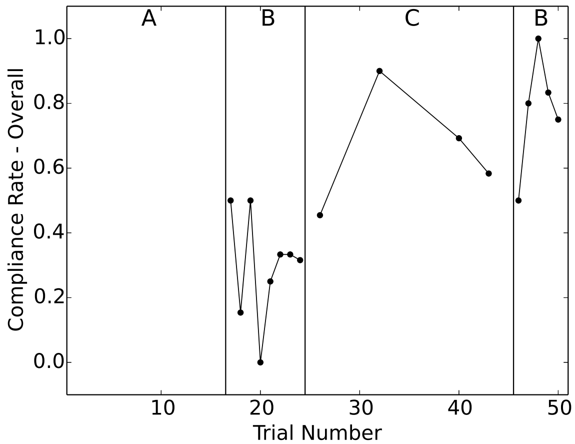
With this in mind, we can attempt to answer what happened in Phase C that resulted in an increase of robot prompts effectiveness in second Phase B. The first observation we can make is by plotting only the level 0 parent involvement trials, where the robot is left alone with the child. The Compliance Rate and Hard Compliance Rate for robot alone trials is shown in Figure 2.9. The Not Affected By Prompt Rate for robot alone trials is shown in Figure 2.10. We see a sharp jump rather than a gradual change in both Compliance Rate and Not Affected By Prompt Rate when we move from first phase B to phase C. This suggests that the cause of effectiveness improvement of robot prompts was probably not a gradual processes such as the child learning hand-washing or getting used to the robot or the washroom environment. Instead, it was a sudden event experienced at the start of Phase C – the joint prompting from the parent telling the child to follow the robot. We also believe that, although the improvement in robot prompt effectiveness was immediate, a long training phase C was needed for the child to retain the behavior. Note that Hard Compliance Rate did not show immediate change in levels. This is possibly due to its low sample size, since hard compliance scenarios (child attempted a step before prompt and then was prompted for another step) were not that abundant.

Point form notes for what else to discuss:

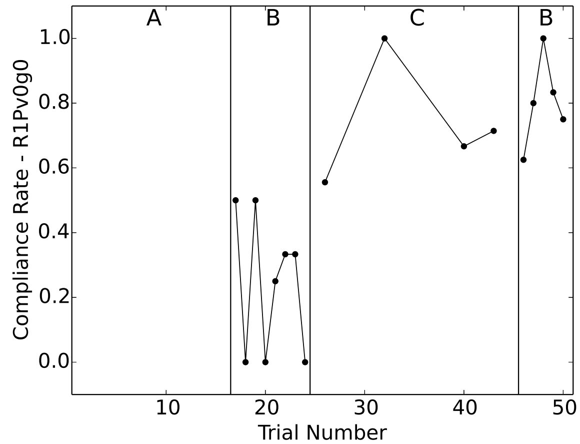
- child engagement

- which step / what circumstance makes child smile / murmur most often? - what does it mean when child smiles / murmurs more?

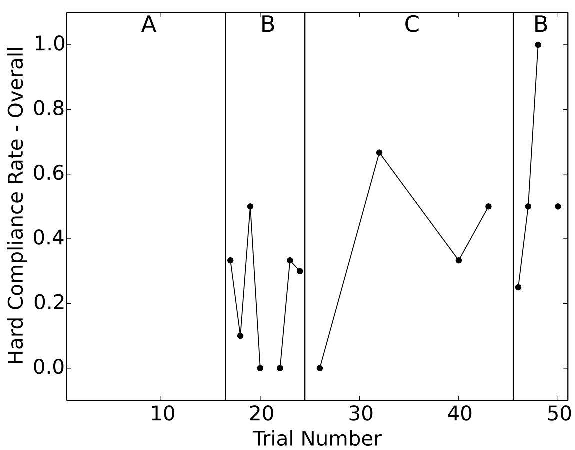
- getting used to the robot / environment: for sure a factor, smiles more later on, and parent also says if it was installed at home and only used 3 times a day, but used everyday, performance may be much better (i.e. fatigue and learning effect)



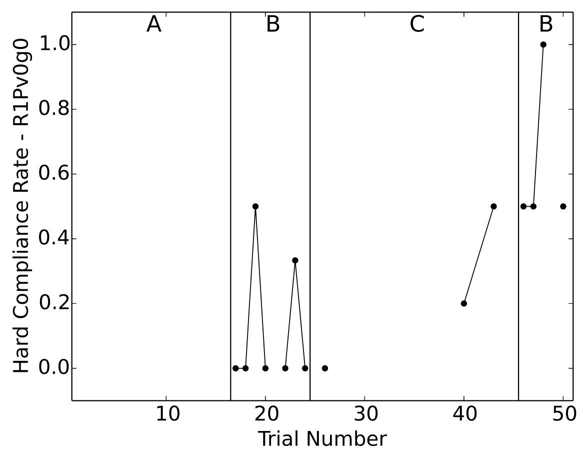
(a) Compliance Rate - Overall - Robot Alone Trials



(b) Compliance Rate - Robot Only Prompts - Robot Alone Trials



(c) Hard Compliance Rate - Overall - Robot Alone Trials



(d) Hard Compliance Rate - Robot Only Prompts - Robot Alone Trials

Figure 2.9: Compliance Rate

- getting used to being prompted

- recognizing robot slowly as an authoritative figure - We assumed that prompts can be ordered by their level of authority over the child, with robot prompts the lowest level of authority, followed by parent non-physical prompts (e.g. verbal prompts and gestures such as pointing and motion demonstrations), and lastly parent physical prompts (e.g. nudging, guiding the arm, and completely executing the step for the child). : what if this assumption is false?

- perceived role of robot - compare with parent - murmurs to parent in protest / asking for permission while looking at parent, but not often to robot

limitation:

- number of subject, sample size of measures

- visual analysis method

- AG not useful? how can we improve? - visual attention not a good indicator of actual attention /

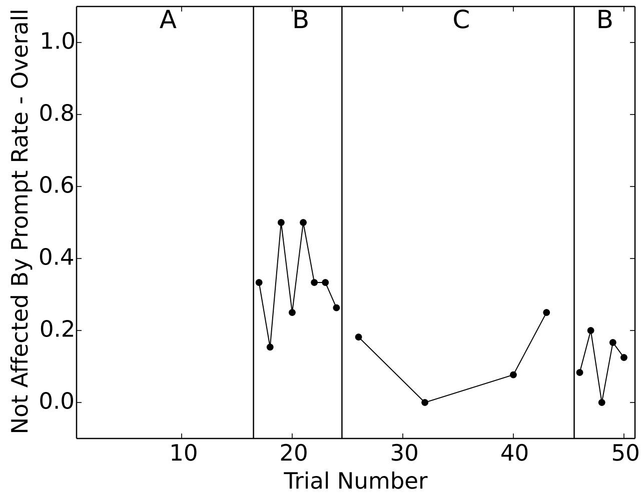


Figure 2.10: Not Affected By Prompt Rate - Robot Alone Trials

engagement

- the explorative nature of study / case-study format of analysis
- inter-rater agreement

Bibliography