Response Prompting for Intelligent Robot Instruction of Students with Intellectual Disabilities

Christopher Reardon[1], Hao Zhang[2], Rachel Wright[3], and Lynne E. Parker[1]

Abstract— Instruction of students with intellectual disability (ID) presents both unique challenges and a compelling opportunity for socially embedded robots to empower an important group in our population. We propose the creation of an autonomous, intelligent robot instructor (IRI) to teach socially valid life skills to students with ID. We present the construction of a complete IRI system for this purpose. Experimental results show the IRI is capable of teaching a non-trivial life skill to students with ID, and participants feel interaction with the IRI is beneficial.

I. INTRODUCTION

In light of significant teacher shortages in the special education field [1], there is a compelling role that socially embedded robots augmenting students' instructional needs could fulfill. Providing instruction to students with intellectual disability (ID) can present unique requirements, and we believe that an intelligent robot instructor (IRI) designed to provide autonomous instruction could meet those challenges.

The term "intellectual disability" refers to students who have an IQ two standard deviations below the mean, generally less than 70 with a mean of 100. Students with ID require a high degree of repetition and face-to-face time with an instructor, as well as consistent, timely, and precise feedback, which a robot could provide tirelessly while avoiding common teaching mistakes.

Socially valid life skills are a group of skills commonly taught to students with ID. Social validity, from the perspective of applied research in special education, refers to the acceptability of the goals, procedures, and outcomes of instruction or treatment [2]. In the context of special education, a *socially valid life skill* refers to knowledge or skills that increase a person's independence of personal, community, or job life. The physical, practical nature of many of these types of skills lend themselves well to robot demonstration and observation. In addition, the methodologies used to teach such skills can be formulated for an IRI.

Prompting is an evidence based practice (EBP) that has been shown to be a successful instructional strategy for teaching discrete and chained tasks [3], and is formulated in a way that can be applied to robotic instructors. Generally, the goal of prompting procedures is to modify human behaviors or teach by assessing the environment and acting to stimulate a targeted behavior. Assistance is provided through prompting in an effort to elicit a desired response. One of the key advantages of prompt response strategies is the possibility of different modalities; common types of prompts are vocal, visual, gestural, models (demonstrations), and physical prompts. Another beneficial aspect of prompting is the ability to "fade" or reduce the intrusiveness of the prompts provided to enable individuals to perform the desired behavior independently. Response prompting has been used with strong success to teach pupils with a wide range of disabilities [3].

In this paper we explore the creation of an autonomous, intelligent robot instructor to teach socially valid life skills to students with ID. We detail the cognitive framework that computationally encodes response prompting as part of a cognitive decision-making process and describe our use of object tracking and interaction through speech and gestures to enable instruction through a common medium between an IRI and a student. The results we present show that our approach is successful in teaching students with ID an important socially valid life skill.

We believe that this is the first use of instruction via response prompting in the cognitive decision making process implemented on an intelligent, autonomous robot.

II. RELATED WORK

Recent works show an emerging interest in applications of robots as teachers. The potential advantages for robot instructors are becoming widely recognized: in [4] social robots interacted with 18-24 month-olds for early childhood education; [5] studied kindergarten-aged children and provided guidelines for using robot tutors in classrooms; the impact of a robot's perceived social role on a student's likeliness to seek educational assistance was examined in [6]; the effects of personalizing a robot's instruction based on an online skill assessment were presented in [7].

A large number of works have examined the effects of embodied robot instructors, coaches, and tutors on outcomes and perceptions. Bainbridge et al. [8] discovered that a physically present robot's commands are more likely to be obeyed than a video representation of the robot, and showed that physical robots are afforded more of the aspects of a human, such as obedience of unusual instructions and physical personal space. Leyzberg et al. [9] showed that instruction from a physically present robot, compared to instruction from a video of the robot, from an audio recording, and no instruction at all, performed best when providing puzzle-solving advice, as measured by puzzle-solving time and self-

[1]Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN 37996, USA `creardon@utk.edu and leparker@utk.edu`
[2]Department of Electrical Engineering and Computer Science, Colorado School of Mines, Golden, CO 80401, USA `hzhang@mines.edu`
[3]Department of Theory and Practice in Teacher Education, University of Tennessee, Knoxville, TN 37996, USA `rwrigh15@utk.edu`
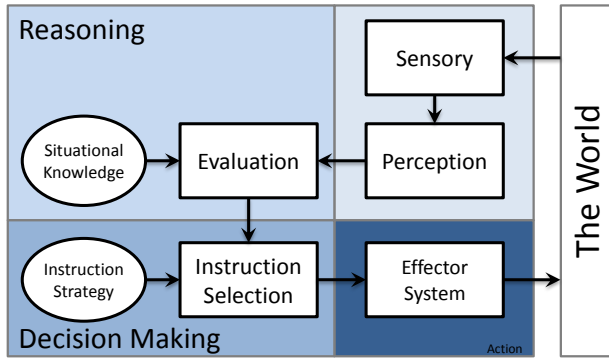
Fig. 1. Cognitive process flowchart. Information flows from lighter to darker shaded process steps.

report measures. Fasola and Mataric [10] found that older adults preferred a physically embodied robot "coach" over a virtual coach for several key social factors, and under robotic coaching performed at a consistently high level. Kidd and Breazeal [11] found a physically present robot is perceived as more enjoyable, credible, and informative than a video character; [12] determined that an embodied robot is more helpful and attentive than a video or simulated character; [13] discovered that cognitively impaired and/or Alzheimer's patients are more engaged by robot treatment than virtual agent treatment; and [14] showed that health advice from a physical robot was followed more often than the same advice from a robot video virtual agent.

Historically much work has been done to create assistive therapy robots for people with disabilities, especially for those with Autism Spectrum Disorder (ASD); surveys of robotics and technological applications to ASD therapy can be found in [15], [16]. Begoli et al. [17] presented a formulation of response prompting as a process ontology for intervention by intelligent agents for children with ASD; however, it was not implemented on a robotic instructor and no experiments were conducted with human subjects. Grezczek et al. [18] presents an approach for feedback in imitation gaming called "graded cueing" that shows promise for encouraging autonomous behaviors in children with ASD that is similar to one type of response prompting.

Our work aims to create an intelligent, autonomous system for instruction, with a particular focus on teaching socially valid life skills to students with ID. Furthermore, because the methodology we have selected has been shown to be widely applicable to students of all capabilities, we believe that the approach should generalize and the lessons learned should be useful to other educational robotic applications.

III. APPROACH

A. Response Prompting for Decision Making

In order to successfully teach, the IRI must be able to make decisions based on observations in order to provide feedback to the student. The decision of which feedback to provide is performed using response prompting as part of a cognitive process.

System of Least Prompts (SLP) is a popular variation of response prompting that is well-validated as an EBP for successfully providing instruction of chained and discrete tasks to students with ID. Further, it is applicable to a wide range of students and has been shown effective at teaching a large variety of skills [3]. We have selected response prompting as the theoretical foundation for our IRI system, and for this work we have computationally encoded SLP for use in an IRI's decision making process.

In the SLP instructional methodology, a hierarchy of prompts is arranged from least to most intrusive. At the least intrusive level, no prompt is used. At the most intrusive level, a *controlling prompt*, i.e., one that assures the task will be correctly performed, is used. The hierarchy of prompts is traversed iteratively to provide more assistance and information as needed. At each iteration, the *target stimulus* is presented with the prompt for the current level (initially, no prompt). A constant amount of time is allowed to elapse before and after each prompt, known as the "response interval." When a correct response is given, it is reinforced, regardless of when it occurs (i.e., at any point in the hierarchy). When an incorrect response is given, the prompt level is escalated. Possible outcomes of each iteration include: unprompted correct, unprompted error, prompted correct, prompted error, and no response error. The goal of SLP is for students to respond correctly before any prompt is delivered, at the lowest level of the hierarchy. SLP is unique in response prompting in that as the student answers correctly at lower prompt levels, a process of "self-fading" occurs, where the student's answers themselves are what determine the rate the intrusiveness of instruction is decreased.

Fig. 1, inspired by [19], illustrates the cognitive framework at a high level. In the cognitive process, an interpretation of the states and actions of the world is created by taking basic sensory information from the world and perceiving information salient to the task at hand. Then, the system reasons on that information, given the knowledge of the task at hand, to generate through evaluation (e.g., about the human activities being observed and the correctness of a response) higher-level representations about the scenario. Using that higher-level information, a decision is then made. In the IRI scenario, this involves using the encoded instructional methodology to select the correct instruction response (e.g., present stimulus, prompt, consequence, reinforcement). Finally, the action is articulated in the world via the navigation and motor system.

The general instruction process is shown in Fig. 2a. The process we have encoded is an adaption for an IRI of the SLP methodology, which leads to a strongly defined process that has been shown to be successful when used by human teachers. The instructional intervention begins with the IRI giving an introduction and general instructions for the scenario. Task Instruction is the formal term for the step of introducing the task and presenting the target stimulus. Next, the IRI selects a prompt. Initially, there is no prompt; that is, the student is given the opportunity to present an answer independently. The student then responds, while the
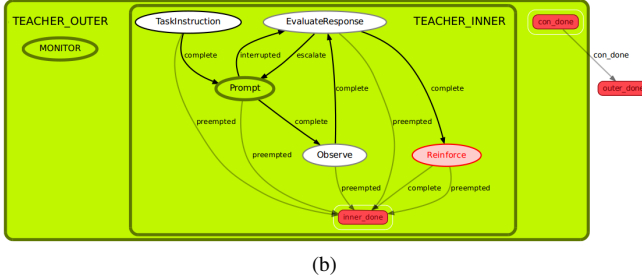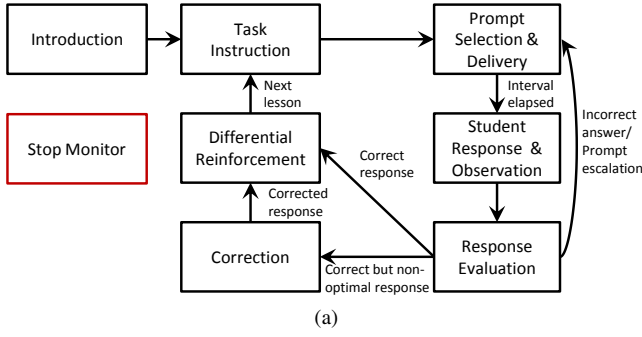
(a)

(b)

Fig. 2. (a) shows the overall instruction process based on response prompting methodology; (b) shows the internal representation generated by ROS SMACH.

Fig. 3. The interaction setting for instruction.

IRI observes the student to determine whether the student is idle, the task is complete, or a period of time, formally known as the response interval, has elapsed. The Response Evaluation determines the result of the step: if an incorrect answer is given, the IRI then uses the evaluated result information as part of the process to select the appropriate response; if a correct answer is given, the IRI provides reinforcement; if a correct but non-optimal answer is given, a correction occurs before reinforcement. All reinforcement is positive. The type of reinforcement is differential, in that it is tailored to the level of intrusiveness of the prompt that was required: students who require less intrusive prompts are rewarded with an increasingly "excited" verbal reaction from the IRI; an independent correct response (i.e., no prompt was required) also triggers a gesture reinforcement (a "thumbs up"). After reinforcement, the student is asked if he or she would like to continue with another lesson. Through repetitions of this process, the goal is for the student to require less and less intrusive prompting, to the point of performing the task independently.

The most important decision making our IRI system performs takes place between the Response Evaluation and the Prompt Selection & Delivery steps. Response evaluation begins when the robot has observed the student to be idle over a specified idle interval, or a longer response interval time has elapsed. Response evaluation determines the type of response: correct, incorrect, partially correct, or no response. That information, coupled with the known information about the previous prompt and human's state (active or inactive), is what the IRI uses to decide the appropriate feedback. In the case where the student is perceived to be actively providing a response during prompt delivery, the IRI immediately reevaluates and presents an updated prompt, if necessary.

Throughout the interaction, a concurrent process is implemented that monitors for a verbal stop command from the student (Fig. 2). Before beginning the experiment, students are informed that any use of the words "stop" or "exit" in combination with the name of the robot, initiates an immediate shutdown of the robot. This provides an additional layer of safety and comfort for the students beyond the supervision of the experiment operator.

B. Object Tracking and Interaction Gestures

Our approach is designed to teach socially valid life skills that require object manipulation and discrimination. In our setting, a human student and an IRI stand across from each other at a table (Fig. 3) and the IRI presents a lesson scenario.

The IRI observes the performance of the requested task, and provides feedback to the student. Observation is achieved through a custom object-tracking system, and interaction and feedback is provided through speech, speech recognition, and gestures.

In an instructional scenario where both a human and robot must interact with objects, the IRI must be able to observe the objects to accurately interpret the performance of the student, make proper decisions, and provide correct instruction; therefore, a fast, accurate method of tracking objects is critical.

Our system uses a high-definition camera mounted under a table with a transparent surface (Fig. 4). To create a simple yet highly accurate and efficient solution to the object tracking problem, known color information about objects being tracked is used. In the event of similarly colored objects, small colored tags are discretely affixed to the bottom of the objects.

Custom object tracking software leveraging OpenCV [20] has been created. Objects are segmented in the image using contours derived from HSV ranges, and positions defined as the contour centroid. Orientations, when applicable, are calculated using methods appropriate for the shape. This approach provides live, highly accurate location and orientation information of the objects on the table surface. Pose information for each object is published into the Robot Operating System (ROS) framework using the tf [21] coordinate frame package.

One challenge for any vision system is accurately tuning it for use, particularly when deployed in different lo-
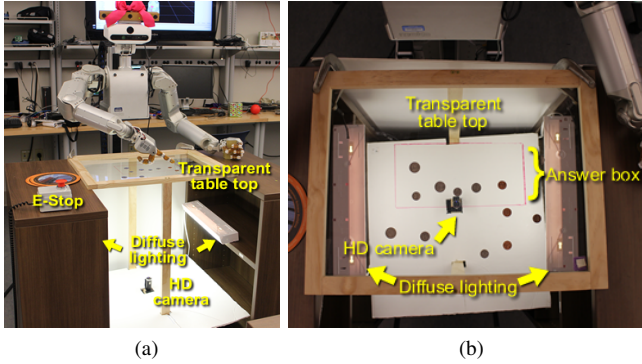
(a)   (b)

Fig. 4. The table setup for the instructional setting from the student view (a) and overhead (b).

Fig. 5. The object tracker GUI, with live, adjustable parameters on top and the annotated live image on the bottom. Right zoom shows an enlarged view of the annotated image. Annotations include position, orientation, size, centroid location, and bounding box for each object.

cations with varying lighting conditions. To address this, a convenient GUI has been created that provides a live, annotated view of extracted location, orientation, size, and identification of objects and allows online adjustment of vision parameters. In the GUI, a live reading of the HSV values at the current cursor location is projected to allow the experiment operator to quickly identify and recalibrate as necessary. Fig. 5 shows a screen capture of the GUI displaying the annotated image and calibration interface.

Measured performance of our object tracking system shows it is both highly accurate and efficient. To compute the accuracy of our object tracking system, we analyze the error distance between detected and ground-truth centroid, as well as a standard object-matching approach using bounding boxes. For matching accuracy we use an approach similar to [22] to evaluate each frame with the PASCAL [23] detection measure by calculating the area of overlap of the ground truth bounding box BB_{gt} and the detected bounding box BB_{dt} of each detected object. The PASCAL measure states that the overlapping area a_o must exceed 50%, as defined in Equation 1, to be considered a match.

$$a_o \doteq \frac{area(BB_{dt} \cap BB_{gt})}{area(BB_{dt} \cup BB_{gt})} > 0.5 \qquad (1)$$

The object tracking accuracy consistently exceeds the PASCAL matching threshold of 50%, with a mean a_o of 0.86 and minimum a_o of 0.6. The mean error distance between detected and ground truth centroids is less than 2mm.

To measure the efficiency of the object tracking system, we record the time until detection of all objects is complete in a live experimental setting for different quantities of coins placed in varying positions throughout the table. Time to detection of 20 objects, which we have selected as the standard quantity of coins for the *making change* task, since it is both twice number of coins sufficient to make change for any price and provides a substantial number of possibilities for students to respond correctly or incorrectly, is 33ms (p=0.95). Using twice that number of objects increases our detection time by 20ms (p=0.95). Combined with total system overhead, our tracking system operates at a frequency above 70Hz.

A second essential need for a competent system for interactive instruction is the ability of the robot instructor

TABLE I

CONFUSION MATRIX FOR GESTURE TEST

		Coin named		
		T	F	
Perceived	T	42 (0.33)	3 (0.02)	PPV:0.93
	F	10 (0.08)	71 (0.56)	NPV:0.88
		TPR:0.81	TNR:0.87	ACC:0.90

to make gestures that are interpretable by, and acceptable to, the human student. Our system uses a collection of physical cues to provide a realistic, attentive appearance, as well as gestures to provide explicit, physical instruction. When the IRI is speaking she uses skeleton tracking to turn her head to face the student, to give the impression of maintaining "eye contact," and when the student is performing a response, the IRI faces the objects being manipulated. When providing instruction, the IRI uses the poses of objects extracted by the object tracking system to point directly at the objects to which she is referring. She is also able to gesture at objects for other procedural purposes, such as asking the student to reset the table environment in between trials. The IRI also uses gestures when providing differential reinforcement.

To examine gesture interpretability, a simple test where the robot interacts with a student with ID was conducted where the robot attempts to deceive the student half of the time by selecting a coin to gesture to, but then either asks whether the coin is the type of coin she is gesturing to, or names a different random coin type, with an equal probability. A confusion matrix is presented in Table I. Over 126 samples, the true positive rate was 0.81, true negative rate 0.87, positive prediction value 0.93, false omission rate 0.12, and overall accuracy 0.90.

Together, the object tracking and gesture components, when combined with out-of-the-box speech and speech recognition, form a complete interaction system.

IV. EXPERIMENTS

A. Robot

The primary robotic hardware for this research is a humanoid robot named Rosie (Fig. 4a). Rosie is a Meka

Robotics M3 mobile humanoid robot with two 7 degree of freedom (DOF) series-elastic arms, two 5-DOF hands, and a sensor head with 2-DOF movement. For this research, she makes use of one PrimeSense short-range camera, one USB camera, a Bluetooth microphone, and stereo speakers.

Rosie is equipped with two desktop-level PCs: one providing real-time functionality of the base, arms, hands, and lift; the second dedicated to the vision and audio components.

All behavior software is implemented in the Robot Operating System (ROS) open-source framework. At a high-level, Rosie's functionality includes collision-aware trajectory-based movement planning using MoveIt! [24] integrated through the ROS action request/service paradigm; speech recognition using CMU PocketSphinx [25]; speech synthesis using eSpeak [26]; and 3D human tracking using NITE [27], all of which are leveraged in this work.

B. Instructional Task

For this research, the task selected for the IRI to teach is the task of *making change*. The ability to make change, i.e., to calculate the correct quantities and denominations of currency that should be exchanged after a cash transaction is made, is a valuable skill for both the workplace and independent living, and therefore fits the definition of a socially valid life skill.

To set up this task, we use the configuration described in Section III-B. We use standard U.S. coins (quarters, dimes, nickels, and pennies) and affix a colored tag to one side of each coin, which is placed facing down, towards the camera.

The general outline of an experimental trial begins with Rosie, our IRI, providing the task instruction and then asking the student to show the correct change for a dollar for a randomly selected price less than $1, which is referred to as the target stimulus. After presenting the target stimulus, the student presents an answer by placing coins into a specially delineated area of the table (an "answer box"). Rosie observes and selects the appropriate feedback at the correct time using the encoded decision making approach.

For the *making change* scenario, a prompt hierarchy has been encoded as a decision tree, and incorporated into the overall cognitive framework. We have defined four levels of increasingly intrusive prompts, with unique interactions for each of the possible response types that could result in a prompt (incorrect, partially correct, and no response). Table II shows an overview of the resulting prompt hierarchy where prompts are arranged from least to most intrusive. *Verbal Cue 1* is the least intrusive; Prompt level *Direction 2* serves as the *controlling prompt*. Responses: NR - No Response, PC - Partially Correct response, I - Incorrect response; PC is a special case where all of the coins in the response are part of the solution, and none are not part of the solution, i.e, the student is progressing towards a solution.

C. Evaluation

1) Single Case Experimental Design: Single case experimental design (SCED) is common in special education research because it allows for the participants to serve as their

TABLE II
PROMPT HIERARCHY FOR THE *making change* TASK USING SLP.

Prompt Lvl.	Resp.	Prompt Description
Verbal Cue 1	NR	Verbal interaction to determine how much change is due
	PC	Verbal encouragement, verbally provide goal
	I	Same as NR
Verbal Cue 2	NR	Verbal interaction to determine which coin to begin with
	PC	Verbal encouragement, verbally provide goal + shortage between current state and goal
	I	Verbal encouragement, provide goal + excess between current state and goal
Direction 1	NR	Gesture to correct first coin, verbally provide goal
	PC	Gesture to correct next coin, verbally provide goal + shortage
	I	Gesture to coin to remove, verbally provide excess
Direction 2	NR	Gesture to each coin to add, wait until added
	PC	Same as NR
	I	Gesture to each coin to remove, wait until removed, then same as NR

own control data for the purpose of comparing performances between at least two experimental phases as opposed to comparison between groups or subjects [28].

A multiple baseline across participants single-subject experimental design was used to determine whether a causal or functional relation exists between the delivery of the independent variable (IV), intelligent robot instruction, and significant increases in the dependent variable (DV), the acquisition and maintenance of the skills required to independently make correct change from purchases under $1.00. SCEDs generally involve repeated, systematic assessment of one or more IVs and DVs over time. Because withdrawal of skill knowledge is not possible, this design was selected in order to allow evaluation of intervention effects by controlling for threats to internal validity and to establish a cause-effect relation. Sequentially introducing the intervention across a minimum of three replications of participants allows for experimental control by eliminating the possibility of any observed change occurring due to extraneous factors (e.g., practice or history effects), thus a causal relation can be established [29].

During baseline, each participant was asked to perform a series of trials related to counting amounts of change back to the researcher for purchases of items under $1.00. No feedback, prompting, or assistance was provided to participants under the baseline condition. At least three trials were conducted until data were considered stable. Stability was defined using the "80%-20%" criteria of the stability envelope, meaning that 80% of the data points fall within 20% of the mean of baseline [28].

Upon verification that participant 1 was unable to correctly perform the skills related to making change independently, and baseline data were considered stable, participant 1 was introduced to the intervention. Participants 2 and 3 continued to be assessed periodically to ensure the skills had not been learned through practice or carry over effects. Once

participant 1 demonstrated an ascending trend of at least three consecutive scores per skill above baseline mean, the intervention phase was introduced to participant 2 while participant 3 remained in baseline. This continued until all three participants had reached acquisition criteria of at least 3 consecutive trials. Section V-A discusses the results of this study.

2) Subjective Acceptance Survey: To evaluate the attitudes of the student volunteers towards learning from a robot, Likert-scale statements and open ended questions were used to collect subjective data. Students were surveyed both prior to and after working with the IRI system. A five-point Likert Scale was used for each statement, and optional open-ended follow-up questions appropriate to each statement (e.g, "Why or why not?", "Please explain") were asked. To ensure a uniform understanding of the questions, surveys were performed orally, with visual aids provided for responses. The pre-assessment survey consists of 18 statements divided into 8 categories, the post-assessment 30 statements in 14 categories, with 2-3 statements each category, and a summative analysis was applied. Section V-B discusses the results of this survey.

V. RESULTS AND DISCUSSION

A. Participant Performance

The task of *making change* is treated as two sub-skills: Skill 1, the ability to identify the correct amount of change for a given price and Skill 2, the ability to provide the correct amount of change using the least amount of coins possible. Rosie provides instruction for both sub-skills, with calculator instruction for Skill 1.

Participants in this study are college-age and attend a post-secondary education program for young adults with intellectual and developmental disabilities. None of the students selected for this study were able to perform any of the steps involved in the *making change* task independently. All participants are female with an IQ between 57 and 67. All three received special education services throughout school under intellectual disability and earned modified high school diplomas. In addition, participant 1 has a dual-diagnosis of emotional disturbance.

To score the participant's performance, points are assigned on a weighted scale based on prompt levels, as in [30], using a 100-point scale. If a student responds 100% independently, a score of 100 points is recorded. Because each solution has n coins, for Skill 2 there are at most n possible controlling prompts. So, for each correct coin placement step performed prior to the delivery of the controlling prompt, a score of $100/n$ points is recorded. In the worst case, where a student does nothing and Rosie directs each coin (Direction 2 in Table II), a score of 0 points is recorded. For each less intrusive prompt, points are deducted in proportion to the intrusiveness: $-50/n$ for prompts regarding the value of the coins, $-25/n$ for general prompts, and $-5/n$ for prompts related to an inefficient combination (e.g., 5 pennies vs. a nickel). For Skill 1, an independent correct response was given 100 points, a correct response with prompting/instruction
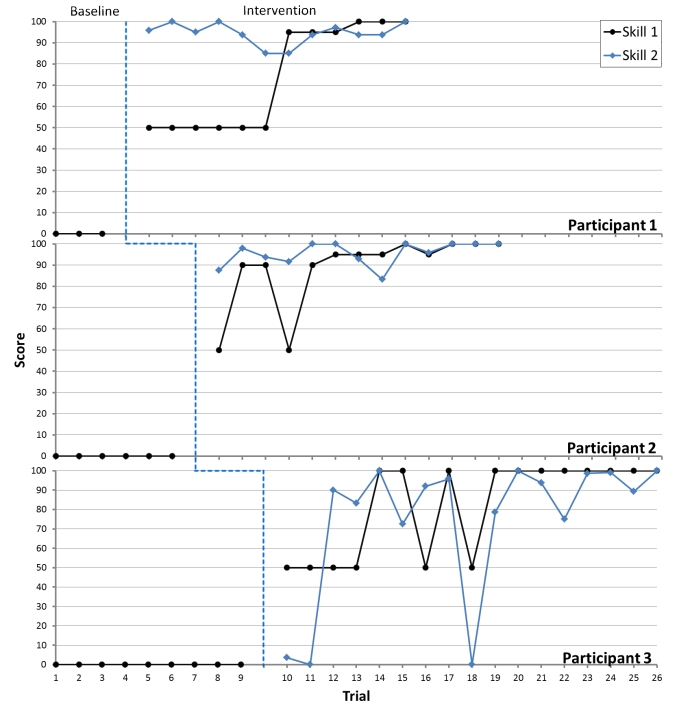

Fig. 6. Results from the *making change* experiment using multiple baseline across participants.

was given 50 points, and a 5 point deduction was applied for each missed response in the calculator instruction sequence. Fig. 6 shows results for each student's performance across baseline, Skill 1 and Skill 2 conditions.

With the IRI the students were able to achieve skill acquisition to mastery, defined as 100% correct performance of Skill 1 following calculator instruction and 100% independence of Skill 2. While data collected during baseline conditions verified that the participants were not able to perform the making change task, specific prerequisite knowledge was not assessed, such as coin value identification and the ability to follow Rosie's directions. Participant 3 demonstrated more limited understanding of these prerequisites than participants 1 and 2, and therefore had a steeper learning curve to achieve criteria for acquisition and mastery. Such considerations will be applied to future studies of robot interventions for people with ID.

B. Subjective Acceptability Survey

An acceptability study was conducted to examine the opinions of students under Rosie's tutelage. Survey responses were scored from -2 (Strongly Disagree) to 2 (Strongly Agree), and averaged across categories. In Table III, we see categories from the pre- and post-instruction survey results. The initial results of the assessment of students' opinions prior to working with a robot instructor showed mixed enthusiasm for the experience; however, post-instruction results show a positive opinion of the overall experience and performance of the robot. Compared to the students' willingness to work with a robot pre-instruction, the students showed greater willingness to work with Rosie again.

TABLE III
ACCEPTANCE SURVEY RESULTS.

Category (Pre-Instruction)	S1	S2	S3
Do you like robots?	1	2	0
Have you been exposed to robots before?	-0.33	0.67	-1
Are robots useful?	1.5	1.5	0.5
Would you learn from a robot?	-0.5	1	1
Category (Post-Instruction)	**S1**	**S2**	**S3**
Was the robot good or bad overall?	1.5	1.5	1.5
Was the robot knowledgeable?	1.5	1	2
Did you trust the robots instructions?	1.5	1.5	0.5
Did you obey the robot?	2	1	1.5
Was the robot easy to learn from?	2	1	-0.5
Would you work with the robot again?	2	1.5	2

The open-ended question answers highlighted this sentiment. One student had several suggestions for additional teaching tasks. Students could also envision the advantages of a robotic instructor, with comments about the ability to "give [human] teachers a break," and the increased ease of learning from a robot, because where teachers move at the pace of the class the "pace could be better" with a robot.

VI. CONCLUSION

For the purpose of teaching socially valid life skills to students with ID, we have presented the first implementation of response prompting for cognitive decision making on an Intelligent Robot Instructor. We have detailed the creation of a functional system to provide essential perception and interaction capabilities for this purpose, with a fast and accurate solution for object recognition, and salient gesturing for interaction. Experiments using SCED methods show that an IRI is capable of teaching students with ID the selected task. We have also presented subjective results from subjects in our volunteer student group that show enthusiasm for this work. We believe that through the advancement of intelligent, autonomous robot instructors there is the potential to empower students with ID, and others, to lead more independent lives.

VII. ONGOING AND FUTURE WORK

The formulation of the response prompting approach and the vision, tracking, and gesture system are being used in ongoing research aimed at examining how a social robot can be used to teach students with ID socially valid life skills.

Planned future directions of this research include examination of other instructional methodologies for cognitive decision making, application to new task environments, and group instruction.

REFERENCES

[1] E. E. Boe, "Teacher demand, supply, and shortage in special education," *Handbook of Research on Special Education Teacher Preparation*, p. 67, 2014.
[2] S. L. Foster and E. J. Mash, "Assessing social validity in clinical treatment research: issues and procedures." *Journal of consulting and clinical psychology*, vol. 67, no. 3, pp. 308–19, 1999.
[3] M. Wolery, M. Ault, and P. Doyle, *Teaching students with moderate to severe disabilities: Use of response prompting strategies*. Longman, 1992.
[4] J. R. Movellan, F. Tanaka, I. R. Fasel, C. Taylor, P. Ruvolo, and M. Eckhardt, "The rubi project: a progress report," in *Proc. Human-Robot Interaction*, 2007.
[5] E. Hyun, H. Yoon, and S. Son, "Relationships between user experiences and children's perceptions of the education robot," in *Proc. Human-Robot Interaction*, 2010.
[6] I. Howley, T. Kanda, K. Hayashi, and C. Rosé, "Effects of social presence and social role on help-seeking and learning," in *Proc. Human-Robot Interaction*, 2014.
[7] D. Leyzberg, S. Spaulding, and B. Scassellati, "Personalizing robot tutors to individuals' learning differences," in *Proc. Human-Robot Interaction*, 2014.
[8] W. A. Bainbridge, J. Hart, E. S. Kim, and B. Scassellati, "The effect of presence on human-robot interaction," in *Proc. Robot and Human Interactive Communication*, 2008.
[9] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati, "The physical presence of a robot tutor increases cognitive learning gains," in *Proc. Cognitive Science*, 2012.
[10] J. Fasola and M. Mataric, "A socially assistive robot exercise coach for the elderly," *Journal of Human-Robot Interaction*, vol. 2, no. 2, pp. 3–32, 2013.
[11] C. D. Kidd and C. Breazeal, "Effect of a robot on user perceptions," in *Proc. Intelligent Robots and Systems*, 2004.
[12] J. Wainer, D. J. Feil-Seifer, D. A. Shell, and M. J. Mataric, "Embodiment and human-robot interaction: A task-based perspective," in *Proc. Robot and Human Interactive Communication*, 2007.
[13] A. Tapus, C. Tapus, and M. Mataric, "The role of physical embodiment of a therapist robot for individuals with cognitive impairments," in *Proc. Robot and Human Interactive Communication*, 2009.
[14] S. Kiesler, A. Powers, S. R. Fussell, and C. Torrey, "Anthropomorphic interactions with a robot and robot-like agent," *Social Cognition*, vol. 26, no. 2, pp. 169–181, 2008.
[15] B. Scassellati, H. Admoni, and M. Mataric, "Robots for use in autism research," *Annual Review of Biomedical Engineering*, vol. 14, pp. 275–294, 2012.
[16] S. V. Wass and K. Porayska-Pomsta, "The uses of cognitive training technologies in the treatment of autism spectrum disorders," *Autism*, p. 1362361313499827, 2013.
[17] E. Begoli, C. L. Ogle, D. F. Cihak, and B. J. MacLennan, "Towards an integrative computational foundation for applied behavior analysis in early autism interventions," in *Proc. Artificial Intelligence in Education*. Springer, 2013, pp. 888–891.
[18] A. A. Jillian Greczek, Edward Kaszubski and M. J. Matarić, "Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders," in *Proc. Robot and Human Interactive Communication*, Edinburgh, Scotland, UK, 2014.
[19] D. Isla, R. Burke, M. Downie, and B. Blumberg, "A layered brain architecture for synthetic creatures," in *International Joint Conferences on Artificial Intelligence*, 2001.
[20] G. Bradski, *Dr. Dobb's Journal of Software Tools*, 2000.
[21] T. Foote, "tf: The transform library," in *Proc. Technologies for Practical Robot Applications*, 2013.
[22] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *Proc. Pattern Analysis and Machine Intelligence*, 2012.
[23] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.
[24] I. A. Sucan and S. Chitta. (2014) MoveIt! http://moveit.ros.org/. [Online]. Available: http://moveit.ros.org/
[25] Carnegie Mellon University. (2014) CMU Sphinx http://cmusphinx.sourceforge.net/. [Online]. Available: http://cmusphinx.sourceforge.net/
[26] eSpeak. (2014) eSpeak text to speech http://espeak.sourceforge.net/. [Online]. Available: http://espeak.sourceforge.net/
[27] OpenNI Consortium. (2014) OpenNI/NITE http://wiki.ros.org/nite. [Online]. Available: http://wiki.ros.org/nite
[28] D. L. Gast and A. D. Spriggs, "Visual analysis of graphic data," *Single subject research methodology in behavioral sciences*, pp. 199–233, 2010.
[29] R. H. Horner, E. G. Carr, J. Halle, G. McGee, S. Odom, and M. Wolery, "The use of single-subject research to identify evidence-based practice in special education," *Exceptional Children*, vol. 71, no. 2, pp. 165–179, 2005.
[30] M. J. Ault and A. K. Griffen, "Teaching with the system of least prompts an easy method for monitoring progress," *TEACHING Exceptional Children*, vol. 45, no. 3, pp. 46–53, 2013.