Pervasive and Mobile Computing 8 (2012) 863–882

Contents lists available at SciVerse ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Pervasive multimedia for autism intervention

S. Venkatesh [a], S. Greenhill [b], D. Phung [a], B. Adams [b,*], T. Duong [a]

[a] Centre for Pattern Recognition and Data Analytics (PRaDA), School of Information Technology, Deakin University, Geelong, Victoria, Australia
[b] Institute for Multi-sensor Processing and Content Analysis (IMPCA), Department of Computing, Curtin University, Perth, WA, Australia

ARTICLE INFO

Article history:
Available online 16 July 2012

Keywords:
Assistive technology
Health
Autism
Therapy
Intervention

ABSTRACT

There is a growing gap between the number of children with autism requiring early intervention and available therapy. We present a portable platform for pervasive delivery of early intervention therapy using multi-touch interfaces and principled ways to deliver stimuli of increasing complexity and adapt to a child's performance. Our implementation weaves Natural Environment Tasks with iPad tasks, facilitating a learning platform that integrates early intervention in the child's daily life. The system's construction of stimulus complexity relative to task is evaluated by therapists, together with field trials for evaluating both the integrity of the instructional design and goal of stimulus presentation and adjustment relative to performance for learning tasks. We show positive results across all our stakeholders – children, parents and therapists. Our results have implications for other early learning fields that require principled ways to construct lessons across skills and adjust stimuli relative to performance.

1. Introduction

"Point to the SMALL dinosaur". Sam points to the correct flash card. Three hours of therapy ends, leaving Sam's mum, Jae, relieved and exhausted. She is too tired to cook dinner from hours spent the previous night searching for dinosaur pictures, enlarging, reducing, printing, laminating and cutting, while Sam rocked himself in a corner. Sam is autistic. He is 2 years old. He does 30 h of therapy every week.

Sam, Jae was told, is one of the 1 in every 150 children who has Autistic Spectrum Disorder (ASD) [1] — a neurological disorder with disruption of early developmental processes. He may show impairment in social interaction, communication, cognitive functioning, and adaptive behaviors. The societal cost over the duration of the life of one ASD individual has been estimated at $3.2 million.[1] This cost can be reduced by as much as 65% if appropriate early intervention is applied. But wait-lists are growing, leaving parents, who lack expertise, to deliver the best-practice therapies, which use flash cards to teach concepts — objects, colours, actions, social situations. For each concept, pictures must be collected, printed, and laminated. Despite this labour, content usually lacks sufficient variety, leading to boredom, over-learning, and an inability to generalise. Moreover, paper-based approaches are unable to teach many concepts effectively — e.g., dynamic concepts such as verbs. Research into computer-based support has targeted early detection [2,3], and affect recognition [4], but has not addressed automated stimulus generation. Proprietary systems, such as *TeachTown*, have limited content, deliver stimuli rigidly, and lack the ability to personalise delivery to the ASD individual, which is so

* Corresponding author.
E-mail addresses: svetha.venkatesh@deakin.edu.au (S. Venkatesh), s.greenhill@curtin.edu.au (S. Greenhill), dinh.phung@deakin.edu.au (D. Phung), b.adams@curtin.edu.au (B. Adams), thi.duong@deakin.edu.au (T. Duong).

[1] http://www.nationalautismcenter.org/.

1574-1192/$ – see front matter
doi:10.1016/j.pmcj.2012.06.010

important to therapy. TeachTown is limited to early cognitive tasks such as sensory and receptive matching. Generalisation is encouraged by off-computer activities through printed activities, but no feedback from these is incorporated into the learning plan. DTT is another system that teaches fundamental concepts of matching, but stops in this early stage of the syllabus.

There is a clear need for deployable assistive technologies that go beyond the matching tasks in cognition to tasks focusing on receptive and expressive language. Receptive and expressive language are the focus of 75% of children in therapy. Systems like TeachTown facilitate receptive language via recorded voice prompts, but lack automated support for the crucial skills of expressive language. Further, although generalisation using offline activities ideally require feedback to be collected and incorporated into the therapy, no systems currently allow this.

We need systems that support parents in therapy delivery in the crucial window of opportunity for their children. These systems must:

– Deliver therapy beyond simple cognitive matching tasks, encode therapist expertise to deliver lesson plans, and perform session logging,
– Use large sets of media stimuli to avoid over-learning, and personalise delivery to the ASD individual and their ongoing performance, and
– Maintain and distribute state so that *therapy delivery can transition seamlessly between therapists (parents, friends, teachers, health professionals)* and *places (home, school, clinic, and the natural environment)*. A pervasive platform would allow therapy to continue anywhere, crucially allowing natural environment tasks to influence formal therapy.

We present *Playpad*, an iPad-based application that implements a ''wait-list program'' for children diagnosed with autism, who face up to a 1 year wait to access an intervention program. In the absence of a trained therapist, our program provides automated multimedia early intervention in this crucial period, teaching basic skills. The syllabus, content and lesson delivery have been designed by trained therapists. *Playpad* leverages a novel framework for delivering generalised, automated, multimedia intervention using platforms that can be used ubiquitously in all places of daily life. Our goal is to take this knowledge-base and syllabus and construct:

– Flexible frameworks for automated stimulus construction, presentation and response recording, for intervention in social and cognitive/visual areas, wherein stimulus adapts to a child's responses and current capacity.
– Intervention for cognitive areas that go beyond current system capability, targeting matching, receptive and expressive language skills.
– Pervasive framework that allows Natural Environment Teaching (NET) to be performed anywhere, with the parent recording their child's responses. NET integrates the physical world with the cyber world, allowing the child to generalise concepts learned on the iPad in the real world.
– Open source infrastructures for multimedia content and metadata, providing reusable content and automatic generation of stimulus from the syllabus, and prompting, reinforcement and stimulus adapted to a child's performance.

We formulate algorithms that deliver these goals, and identify the computational requirements for learning controlled by stimulus, prompting and reinforcement, each of which is carefully crafted to adjust to performance. Playpad records the state of learning within the syllabus for each child, which enables the child to re-enter the learning environment anywhere the iPad can be taken. This ability to provide untethered therapy pervasively makes the framework different to existing commercial systems. Our implementation of NET allows both fully offline and mixed cyber-physical world interactions, such that in both cases responses are recorded. The pervasive nature of the learning platform integrates early intervention into the child's daily life.

To show the usefulness of the framework, we performed three sets of trials and document results from all stakeholders; therapists, children and parents. The trials were performed through our partner *Autism West*. The first trial was conducted with 7 autistic children 4–10 years old, and we verify the instructional framework's ability to adjust stimulus to performance. The second trial was comprehensive; 8 autistic children took part in a two week trial with their parents. We evaluate matching, expressive language tasks and NET. Our results show the framework obtains all of our goals in terms of stimulus presentation and adjustment, thus facilitating learning. Additionally, 2 therapists evaluate the usability of Playpad, which results in unanimous support for all features. This evaluation is independently supported by the parents who participated in the trials. Our third trial involved 16 children over 4 weeks, where therapy was incorporated into participants' daily activities at home. Our results show noticeable progress at skills not previously learnt.

The novelty of this work lies in our construction of generalised learning frameworks for early intervention, a foundation that allows us to go beyond simple matching to complex receptive and expressive language tasks. Our pervasive implementation of NET activities is unique, allowing the child to use the iPad in physical or mixed cyber-physical worlds, while facilitating efficient recording of responses. This feedback is used effectively in therapy.

The significance of successfully deploying this technology is proportional to the massive societal cost of ASD in terms of parental anxiety, stress, and depression [5]. The potential scope of the framework is the much larger context of generic learning.

2. Related background

2.1. Software and assistive devices for ASD

Engineered aids for ASD have targeted three main areas: early detection and measurement of progress; assistive tools for affect recognition; and social skills improvement. Kientz et al. have experimented with toys with embedded accelerometers and wireless sensors to record a child's interactions to improve early diagnosis of ASD [3]. Children with ASD struggle with affect recognition, which led Picard and her team to construct methods for facial recognition, with emphasis on affect, and deploy them on mobile platforms as assistive tools for children [4]. Limited work has been done aimed at teaching social skills in 3D environments that are surrogates for the real world in teaching social skills, but concrete results are yet to emerge [6]. Stanton et al. also addressed the problem of social skill improvement for children with ASD, with early studies using robotic toys [7].

Switching our focus to visual supports, there is ample evidence for the efficacy of systems in the early treatment of mental disorders including autism (e.g., [8,9]). Visual support strategies have been employed through proprietary computer software, such as DTTrainer[2]; pictorial cues, such as the Picture Exchange Communication System (PECS) [10]; and visual social stories, which are concrete, idiosyncratic video narratives used to teach social skills [11]. In particular, computer-assisted intervention has been demonstrated to be effective in teaching language, reducing inappropriate verbalisation, increasing functional communication, and improving generalisation to a child's daily environment [12,13]. Teachtown[3] is another proprietary educational tool that targets autistic children, but suffers from a restricted set of stimuli and adaptation to response. Further, being a closed system, no individualised content can be contributed and used. This lack of flexibility of stimuli construction and content are drawbacks of proprietary systems in general.

Anecdotal stories are emerging on how the iPad is revolutionising learning for autistic children.[4] Recent iPad applications target specific skill shortages. Examples include: Proloquo2Go, iComm, and TapToTalk, which construct an alternative communication system using symbols and text-to-voices; Grace, which helps autistic children to build sentences from images; iCommunicate, which allows storyboarding through pictures and images; First-Then-Visual-Schedule, which enables daily schedules to be taught; iConverse, which translates PECS into a iPad application; AutismExpress, which helps a autistic child understand emotions; and stories2learn, which helps create personalised social stories to teach autistic children about complex social situations. Thus, although these iPad applications target specific shortcomings, there are no systems that target generic, automated stimulus complexity, and frameworks for delivery and recording in the systematic way proposed here.

2.2. Early intervention for ASD

Of the established ASD treatments, Applied Behavior Analysis (ABA) [14] and Discrete Trial Training (DTT) dominate (US National Standards Report on Autism, 2009). These approaches aim to improve cognitive functions such as object labelling and categorisation. Currently, intervention is designed and partially administered by therapists, but almost half of the therapy requires parental involvement. Being paper-based, ABA places a heavy burden on therapists and parents to prepare and deliver a large amount of material, and makes teaching of dynamic concepts, such as verbs, especially challenging. While it is true that Web 2.0 offers a surfeit of accessible media content, it is typically unsuitable for therapy as found. For example, the results of image search often contain too many distracting elements to be directly used, and must therefore be edited. Moreover, record keeping during sessions is vital to ABA and DTT, but opportunity and financial costs mean it is poorly implemented in practice.

ABA's well structured progression formulae, and requirement for record-keeping, lends itself to computational delivery. But we hasten to add that automated delivery is not intended to entirely replace human-delivered therapy. ABA forms the basis of the program delivered on Playpad by means of the underlying content delivery framework described in the following section.

3. A framework for ASD intervention delivery

3.1. Design considerations

The efficacy of Playpad hinges on the creation of a framework for delivering stimuli on Playpad and in physical environments, while adapting to a child's performance. Four aspects are crucial:

– *Stimulus delivery on iPad:* Playpad must automatically construct trials consisting of various stimuli of varying complexity, and adjust lessons and responses based on child performance,

2 www.dttrainer.com.
3 www.teachtown.com.
4 http://www.sfweekly.com/2010-08-11/news/ihelp-for-autism/%20.

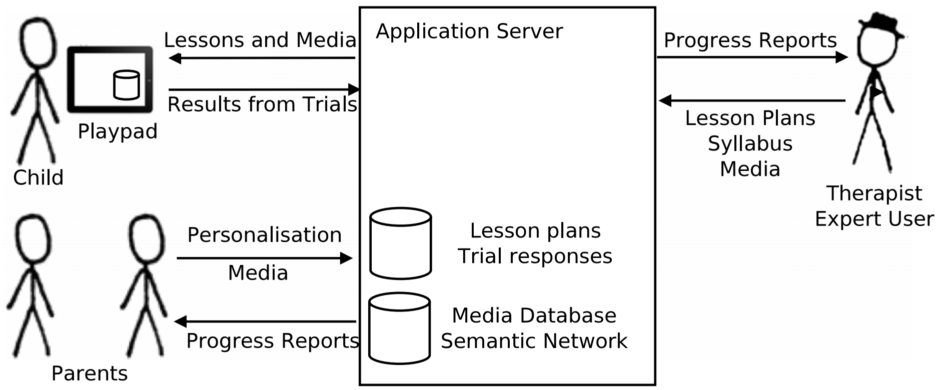


Fig. 1. System architecture.

– *Stimulus delivery in mixed environments:* Playpad must allow for learning in mixed physical-cyber environments. This involves two task paradigms:
  . Playpad presents a stimulus on the iPad, the child responds verbally, the parent records the response on the iPad;
  . Playpad presents a task to perform in daily life that generalises concepts learned on the iPad, the child interacts with the physical world, and the parent records the response on the iPad.
– *Construction of appropriate stimuli:* Playpad must be able to access media content that is annotated with linguistic concepts and relations relevant to behavioral intervention, and do so via a scalable, open system, such that a variety of different learning tasks can be built using shared, user-contributed media and meta-data.
– *Construction of language independent stimuli:* Playpad must be able to provide stimulus and prompts in different languages using a unified framework.

The above requirements fall directly out of the role Playpad assumes or supports: the therapist. Trial delivery must be adaptive because the pattern of development, and progression of those skills, across the ASD population is diverse. User contribution of media, aside from helping achieve a critical mass of stimuli that helps avoid over-learning, addresses another key challenge: engaging children. Children with ASD often have particular domains for which they have a strong affinity, and personalising stimuli to use favourites, where possible, aids engagement.

Playpad users may be as young as 2 years, which necessitates a platform that supports natural interactions such as pointing and touching. It should also be cheap, robust, and portable, in order to be used pervasively within the home and other places of daily life. It must be able to present visual stimulus with "touchable virtual cards" that can be static or moving, and provide prompting, reinforcement, and reward using sound and motion. Not only should lesson preparation time be reduced, but progress recorded automatically for review, in order to further reduce parent/therapist burden, and increase time-on-task. Both on and off Playpad activities must be facilitated.

3.2. System architecture

Fig. 1 shows an overview of the system. The mobile Playpad receives lessons and media from, and reports session results to, the application server, which consists of two components: (i) a database populated with multimedia content, and (ii) a framework for constructing a stimulus set dictated by a syllabus.

The application server resides on the web, accessible via standard web protocols, and hosts a database of media and information about semantic relations. Autism experts design a syllabus which is subsequently translated into a series of lessons a child works through at their own rate. Each lesson consists of media stimuli, which are in part chosen using the encoded semantic relations, and are downloaded as required by Playpad so that lessons can be performed off-line if required. Results from training are fed back into the system, which are used to adapt the lessons as required, and provide progress reports to parents and therapists. Once synchronised with the server, a user's state is available on other synchronised Playpads, allowing the child to move between different devices (e.g., at home, at a therapy provider, at school or kindergarten).

To enable flexible and personalised programs, the multimedia content and metadata is powered by a semantic network. Links in the semantic net facilitate automated generation of stimulus of a specified complexity for intervention tasks. The framework is agnostic to modality and associates any multimedia content with concepts. Additionally, it is language independent, allowing delivery in different languages using the same metadata knowledge-base.

Playpad is software for delivering therapy activities and collecting progress results. An initial version of the system (2009) used a custom built multi-touch surface, but for this study we use iPads. The key requirement is for a robust portable multi-touch display of moderate size, which is increasingly satisfied by a variety of commodity hardware. The client application is implemented in Javascript and HTML, but can be deployed as a native application, allowing data and media to be cached locally on the device so that it can be used off-line. Normally Playpad is used solo, but occasionally two users may need to collaborate on simple partner tasks.

While each stimulus set is presented to the child by Playpad, the child's response may be external to the iPad (e.g. an utterance). Regardless, all responses must be recorded. When stimulus is external to the iPad, the parent records the response on the iPad. When the child performs a Natural Environment task (NET), the parent records the response on the iPad, and this is used to assess how well the child has generalised the learned concepts.

3.3. Populating the database

Given a set of language concepts from the syllabus, we need to identify corresponding example images. Simply using a web image search will not necessarily return suitable images, so it is important that the exemplars be hand-selected. We gather images from the web and other sources, and build a knowledge-base to augment the media collection, consisting of:

- *Concepts*: important to early language learning. This set includes objects (eg. spoon, dog, car), attributes (e.g., red, colour, small) and actions (e.g., walking, eating).
- *Words*: each concept is associated with a set of words that provide textual representations in different languages.
- *Relations*: a set of binary relations between concepts used to represent both taxonomy (e.g., red "is_a" colour, dog "is_a" animal) and functional association (e.g., spoon "used_for" eating).
- *Media*: image, video, and sound files that illustrate concepts. E.g., a picture of a dog, the sound of a dog barking, the sound of the word "dog" spoken in English.
- *Tags*: a set of relations between Media and Concepts. Importantly, tags relate media to abstract concepts, not to words in any particular language. This makes the knowledge representation independent of any particular language. Concepts and relations are translated to the user's chosen language where possible.
- *Utterances*: a set of utterances of a word being spoken by a particular *Voice* (unique for a given speaker and language).
- *Users*: own media and tags, which allows content to be personalised. E.g., "father" is associated with a photo of the user's father.

Most of the semantic information (Concepts, Relations) is built into the system, but it is envisaged other parts (Media, Tags, Words, Utterances) will be added by users. The database is implemented as a web application, which allows files to be uploaded and tagged in a similar manner to existing media sharing systems. Tags can be either canonical concepts or free text. When a new textual tag is entered by a user, the system searches for the corresponding concepts in Words for the user's chosen language. Tags that cannot be resolved to a concept are highlighted and saved as free text, meaning that the annotated items are findable via text search, but not as concept exemplars. Alternatively, tags may be chosen using a tree-like browser to navigate taxonomic or functional relations.

When adding items (Media, Tags, etc.) the user has the option to declare them private or public. Certain concepts (e.g., home, mother, father) are personal to the user and are always returned preferentially. As per best practice of media sharing applications, items are moderated by administrators before becoming publicly accessible.

Certain parts of the system can be extended by expert users, including Words, Utterances and Voices (which are used to translate content into different languages). The semantic network is constrained to a few hundred concepts crucial to early language learning. These can be largely translated in isolation since the system does not need the grammatical rules required to build sentences. In the few situations where verbal instructions must be generated procedurally (e.g., "find what says 'meow'"), this can be done by concatenating fragments of recorded sound. Sentence prompts related to concepts (e.g., "find the car") are pre-recorded so that the sound is natural to the learner. In addition, a small number of motivational phrases are recorded (e.g., "well done", "no, try again").

4. Learning theory: fundamental tasks and generation of stimulus sets

Key stimuli and concepts are required for human development within and across cultures. These include linguistic and cognitive concepts and are largely identified using developmental norms and sequences and developmental theory. A set of first words by age 4 have been identified, consisting of 300+ nouns, verbs, and adjectives. The daily lesson plan is generated from a syllabus and Playpad generates one activity from the following skill set:

- *Imitation:* Requires the child to imitate motor actions;
- *Matching:* Playpad instructs the child to "Find same" and includes two tasks;
  - . *Sensory matching*: for basic auditory and visual (colour and shape) concepts;
  - . *Identity matching*: for objects of a given category, such as nouns and verbs.
- *Receptive language*: Requires the child to respond *receptively* to language. Playpad says the name of an object, and the child touches the target picture on Playpad.
- *Expressive language:* Requires the child to use language *expressively*. Playpad shows an object, and the child names it.
- *Natural environment task:* Playpad suggests task-coordinated activities to be performed by the child and parent in the real world.

Playpad works sequentially through this list for matching, receptive, and expressive tasks. Stimuli are retrieved from the media database by querying for the required items using the relevant sets of tags. Each skill set is examined in more detail below.

4.1. Imitating gross motor skills

Playpad's syllabus includes a list of imitation tasks designed to teach and reinforce basic motor skills. These are categorised in increasing order of difficulty as: Imitation of gross motor skills (e.g., tap table, clap hands); Imitation of actions with objects (e.g., ring bell); Imitation of fine motor skills (e.g., clasp hands, open and close hands); and Imitation of oral movements (e.g., stick tongue out, smile).

Playpad works through this list in order of skill complexity, and either the parent is requested to act the skill, or the action is demonstrated by a video. The child is asked to imitate the skill, and the parent indicates their performance on Playpad via a feedback screen [Yes/No/Prompt].

4.2. Matching

Sensory matching tasks train the child to discriminate visual attributes (such as colour, shape, texture, or objects) and auditory attributes (environmental sounds, vocal sounds, word sounds, or syllables). Matching is done without reference to language or names. For visual matching tasks, the stimuli (target, variants and distractors) are presented on the screen and the child responds by touching one of the images. The child must simply match things that are similar. For these tasks there is only one instruction, "find same". Matching tasks typically present a number of stimuli, one of which is the "target". The child may be prompted for the target using visual or auditory cues.

The learner needs to be systematically shown how to attend to relevant sensory dimensions of target stimuli and ignore those that are irrelevant. Lesson plans specify sets of stimulus concepts (e.g., ["red", "green"] for colour) and sets of dimensions of variation. In many cases, multiple categories are taught simultaneously, so one stimulus concept serves as a distractor for other stimulus concepts in the same set. Normally, therapy starts with few or no variations in irrelevant dimensions, but increases such variation as the simpler stages are mastered. Variation of a stimulus along an *irrelevant dimension* (relative to target concept) produces a *variant*; variation of a stimulus along the *relevant dimension* (relative to target concept) produces a *distractor*. For example, when teaching the concept "colour", colour is the relevant dimension, and a "white circle" is a variant of a "red square", whereas a "red car" is a distractor.

Certain learning activities rely on relationships between concepts. For example, the "What is it used for?" game is a matching task in which the user must identify the correct activity related to an object (or vice versa). Here, taxonomy and multiple functional relations allow the system to easily generate distractors that are "close" to a target concept.

To illustrate, we demonstrate the relation "Spoon is used for Eating". The system generates distractors in this way:

(i) First it identifies "sibling" concepts, by going one level up the concept tree and then back down *using the same relation*. Thus, if a spoon is found in a kitchen, it will choose other objects also found in a kitchen. This is done for each relation that applies to the target, "Spoon".

(ii) Next it identifies objects that share a similar functional relation. E.g., if a spoon is "used for" eating, it looks for other objects that are "used for" something (i.e., other than eating). This process usually returns concepts of the correct type (places, tools, etc.) depending on the relation.

(iii) As a fallback, if nothing can be found via (1) and (2), it includes all concepts for which media items are defined. This sometimes produces nonsensical results (e.g., "Triangle used for Eating"), but this is rare and not necessarily a problem.

4.2.1. Receptive language

Receptive language is the skill of understanding words spoken by others. Again the skill is organised in increasing levels of difficulty as: Follows one step instructions (e.g., sit down, stand up); and Identifies (e.g., body parts, objects, verbs in pictures, familiar people).

One step instructions are presented in the same way as Gross motor skills, above. For identity matching, Playpad presents pictures of objects from the categories, and prompts the child to identify the correct object, e.g., "Find apple", "Find hand". Of the presented images or sounds, one is the target. The child must touch the target, and Playpad records the response.

4.2.2. Expressive language

Expressive language is the skill of producing words to be understood by others. Again the skill is organised in increasing levels of difficulty as: Imitates sounds and words; Labels body parts, household items, familiar objects; Makes a choice; Greets; and Answers social questions.

Playpad presents the image of an object from the named categories, and the child is required to verbally label it. Prompting here is supplied by the parent, external to the iPad. Responses are recorded on the iPad.

4.2.3. Natural Environments Teaching (NET)

Natural environment tasks are crucial for concept generalisation. NET activities are organised for each concept or skill being taught, and structured in three categories: Personal; Play; and Adaptive routines. The idea behind NET is that depending on the skill being learnt, Playpad will suggest activities to be incorporated into play or daily routines.

For example, if ''Matching food'' is being taught, a Play routine will be ''What's in the Bag'' or ''Lets go shopping'' (see Section 6.3 for a descriptions). The parent is given instructions on how to perform the task and what to look for.

Natural Environment tasks are structured as follows:

– *Prepare the parent:* with knowledge of why the skill is important, done through text and video recordings of therapists. Videos are presented as a model to copy from.
– *Practise and feedback:* parents complete the activity. Playpad presents a list of key elements to look for in an activity (e.g., pointing) and provides a screen for recording the child's response.

Two versions of NET are implemented:

– The activity is performed in the real world, and responses are recorded by the parent on the iPad.
– The activity is performed in the real world using a list provided by Playpad, responses are recorded on the iPad, and prompting is done by Playpad and/or parent.

5. Adapting stimulus complexity to performance

In the previous section we detailed the core competencies our framework addresses, and the rationale behind the stimuli used to learn them. But learning is dynamic, and assistive technologies need to be adaptive. This advancement of stimulus complexity is an interplay of stimulus sets, child performance, prompting and reinforcement. Below we detail the algorithms that support this process.

5.1. Measurement

Ongoing measurement is a necessary condition for adapting stimuli to best fit with a child's needs. A record must be kept of what the child is currently doing, what skills have been learnt and how they were learnt, together with those skills and stimuli causing difficulty. This data is used as feedback to adjust the stimulus, guide teaching strategies, and target learning to the developmental needs specific to each learner.

A single response to a stimulus is termed a trial. A session typically consists of 10 trials, which use the prompting strategy presented below (Section 5.4). Mastery within a session is deemed to have been achieved if the accuracy is above 80%. Therapy parameters, here and below, have been provided by therapists, and are derived from best-practice in psychology.

The following measures are also recorded for future analysis and lesson adjustment: Retention — number of correct responses to stimulus without any distractor; Distraction — number of correct responses to a stimulus with distractor being used; Generalisation — number of correct responses in locations other than the original learning context.

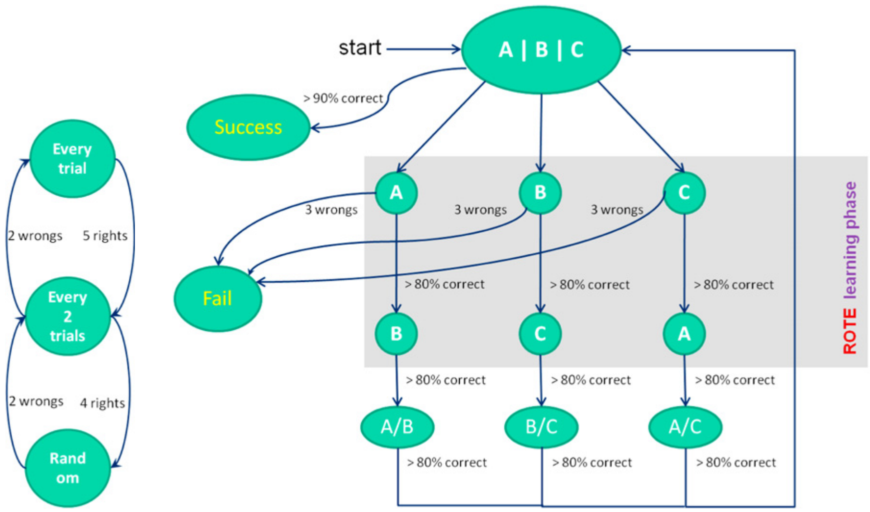
On iPad–Playpad continuously records the results of each trial (stimuli, target, response, correct/incorrect, reaction time) to a local database which is then uploaded to the server for analysis. Accuracy is measured continuously during trials and controls progress through the different stages of a lesson, according to the specified mastery criteria. Normal progress requires correct responses under increasing levels of distraction. If this does not occur, the system falls back to simpler stimuli without irrelevant dimensions of variation, and finally to basic rote learning (as seen in Fig. 2(b)). Once a task has been completed using training examples, we test for generalisation by repeating the task with previously unseen stimuli.

Off iPad–The parent records the responses for off iPad activities. For expressive language training, for each trial the parent can record the responses correct, incorrect, or prompted. For NET activities, the parent can also record how the child performed the task (correct, incorrect, prompted) for each trial.

5.2. Task state and stimulus complexity

For Playpad solo activities every trial response is recorded, and, as noted, for an activity of 10 trials the child must achieve an 80% success rate to progress to the next level of the curriculum. Tasks are separated into numerous steps but broadly into 2 main types of exposure to the learning context: Rote Teaching and Multiple Exemplar Teaching (MET). MET involves presenting a specified number of stimuli at random so the child learns the skill across stimuli and thus generically. For example, if we were teaching receptive labelling we would ask the child to point or touch a given stimulus. If we were doing MET we would rotate between three stimuli (e.g., cat/dog/rabbit) and gesture to the correct one if they made an error. If the child learned all 3 labels within 10–20 trials we would conclude they are effectively learning receptive labelling from gesture. If they failed to learn the labels this way, we would simplify the task to the rote level. The rote level requires that initially only one stimulus and its label be taught at a time (e.g., cat), with additional stimuli added slowly until the learner is rotating between three levels. The advantage of rote learning is that, in learning to discriminate between the stimuli needed (in this case the picture and the word), the child is able to use recent learning history for that stimulus (i.e., for the last trial it was 'cat'; it will be same in this trial), whereas in pure MET the child must discriminate between 3 names and pictures without recourse to recent history for a given stimuli. In short, in MET they must learn receptive labelling as a skill, whereas in rote matching they can learn each stimulus one at a time.

While the goal is to avoid rote because it is slow and developmentally not as efficient, some learners require this level of explicitness in order to understand what is expected of them. Accordingly, Playpad starts every learner at the MET level, and



(a) Reinforcement prompting transition.

(b) State diagram for stimulus set adaptation on equivalence class training for three concepts: A, B and C.

Fig. 2. State transition diagram for reinforcement prompting and stimulus set adaptation.

if they succeed they are moved through the sequence until mastery is established. If they fail, the system takes them to the rote level and attempts to slowly build them up to the preferred MET level. Fig. 2(b) illustrates stimulus advancement relative to performance. It starts in MET with 3 stimuli (A/B/C), before dropping down to rote if three consecutive errors are made at the highest level of prompting (see Section 5.4). The first phase of rote training is to simplify to a single stimulus category (A), followed by a second category (B). If the child achieves 80% accuracy over 10 trials, the task proceeds to simplified MET, which uses pairwise stimuli (A/B, B/C, or A/C). If 80% accuracy over 10 trials is achieved at this level, full MET (A/B/C) is employed.

5.3. Reinforcement

Motivation systems are essential to intervention. Playpad indicates the correct answer visually, by superimposing a flashing star over the correct response, or by audibly saying "good job". Children are awarded a token for correct answers, in this case a picture of a gold coin.

Two categories of reinforcement are implemented in addition to this basic method:

Within a trial — screen based: The system provides exciting sensory feedback on the screen, such as small firework explosions. This kind of reinforcement is provided automatically based on the child's performance, according to the algorithm shown in Fig. 2(a). The system begins at the highest level of reinforcement, which occurs at every trial, and subsequently drops to every other trial if 5 correct answers in a row are given, and finally shifts to random reinforcement if a further 4 correct answers in a row are given. If errors are made, the system shifts back into states that provide more reinforcement, as shown.

Across trials: Once an activity is completed, or a certain number of tokens are acquired, the child is able to access more powerful reinforcers. The child chooses *before* the session to receive either an off-screen reward, such as chocolate, or a screen-based reward, such as a small game involving making ripples, popping bubbles, or painting.

5.4. Prompting

Prompting here refers to indications to a child of the correct response in a given trial. Prompting is crucial for learning, and should be increased or decreased relative to performance. We introduce three levels of prompting, in ascending order: *Level* 0: no prompt is given; *Level* 1: the target pulses; *Level* 2: the target is enclosed in a circle and pulses; and *Level* 3: the target is enclosed in a circle, pulses, and is indicated by a pointing finger.

The state machine that implements the adjustment of prompts for a given teaching concept relative to performance is shown in Fig. 3. In this finite state machine, emitting symbols are of two types: $A{+}{+}$ for two consecutive right answers, and $A{-}{-}$ for two consecutive wrong answers. Any state, if neither A++ nor A−− is achieved, will self-transition unless the total number of trials exceeds 10, in which case it transitions to the terminating state and moves on to the next teaching concept. In such scenarios, the teaching concept has *not* been achieved, and will be recorded by the system for later suggestion in lesson planning. If the terminating state is reached in less than 10 trials, the teaching concept is said to be achieved.

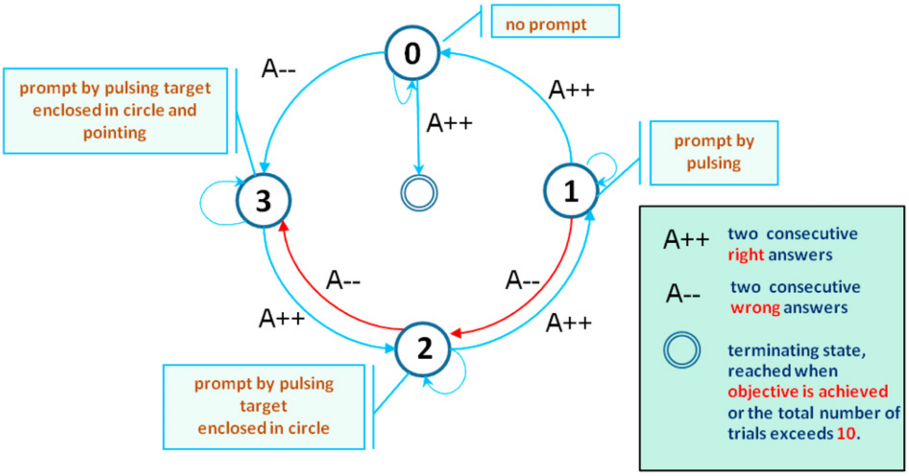


Fig. 3. Finite state machine for prompt adaptation. It emits two types of symbols: A++ for two consecutive right answers, and A−− for two consecutive wrong answers. The degree of prompting is increased or decreased based on a child's performance. iPad tasks begin in State 3; NET tasks begin in State 0. Not shown here is the trial count, which forces a transition to the terminating state when it exceeds 10.

Otherwise, as can be seen in Fig. 3, the system increases or decreases prompt levels based on the number of errors made, while concurrently recording the mode of successes and failures for later use in lesson planning.

Different starting states are used for iPad tasks and NET. iPad tasks start in State 3 to achieve what is termed *errorless learning* through most-to-least prompting. This approach is based on research on stimulus control transfer, e.g., [15]. In contrast, NET tasks start in State 0 thus manifesting least-to-most prompting. The reason for this difference is NET's emphasis on generalisation of skills to the natural environment. While most-to-least is more accurate, it is also slower [16,17].

6. Delivery of daily lessons

We began with an introductory curriculum designed by an autism expert. This includes several types of activity:

- *Solo*: activities implemented on the iPad which a child can perform unassisted. Sensory matching tasks are presented in this way. Stimuli are adjusted to the child's performance, as detailed in earlier sections.
- *Partner*: activities in which child and parent interact via the iPad. Expressive language training is an example. The iPad presents a stimulus, and the child responds verbally. The child may need to be prompted by the parent. The parent records via the iPad whether the child responded correctly or incorrectly, and if prompting was required.
- *Natural Environment Training* (NET): activities that the child and parent can perform during their daily routine. For example, while shopping or preparing a meal, the concepts taught on the iPad are reinforced with daily tasks. The parent records responses via the iPad. NET is crucial to learning generalisation, and we implement both delivery and recording of responses.
- *Reinforcement*: Activities that the child performs with or without using the iPad.

Playpad presents a daily lesson plan (see Fig. 4), allowing the parent to choose from a selection of items that guide the child through the curriculum. iPad activities are delivered via the framework described here. The other activities (NET and Games) are integrated into the system via instructional videos and feedback screens which can be accessed from the lesson plan. Below we detail Playpad's instantiations of the theoretical tasks introduced in Section 4.

6.1. Matching tasks

Visual sensory matching: a row of three images is shown, and the child must identify which one matches the stimulus. This is done using only sensory cues, and does not require the images to be understood (e.g., see Fig. 5). In the example, the target is the cup, and a visual prompt is shown for the correct stimulus.

Auditory sensory matching: a target sound and distractor are represented as a row of two buttons. The cue is another button on a separate row. The system plays a sound whenever a button is touched, and the user must match the target to the cue.

Symbolic processing — receptive labelling (word): similar to sensory visual matching, except that the "cue" is not an image, but the word for the target (e.g., "find the drum").

Visual matching — receptive auditory (sound): similar to Receptive Labelling (Word), except that the prompt is the sound an object makes.

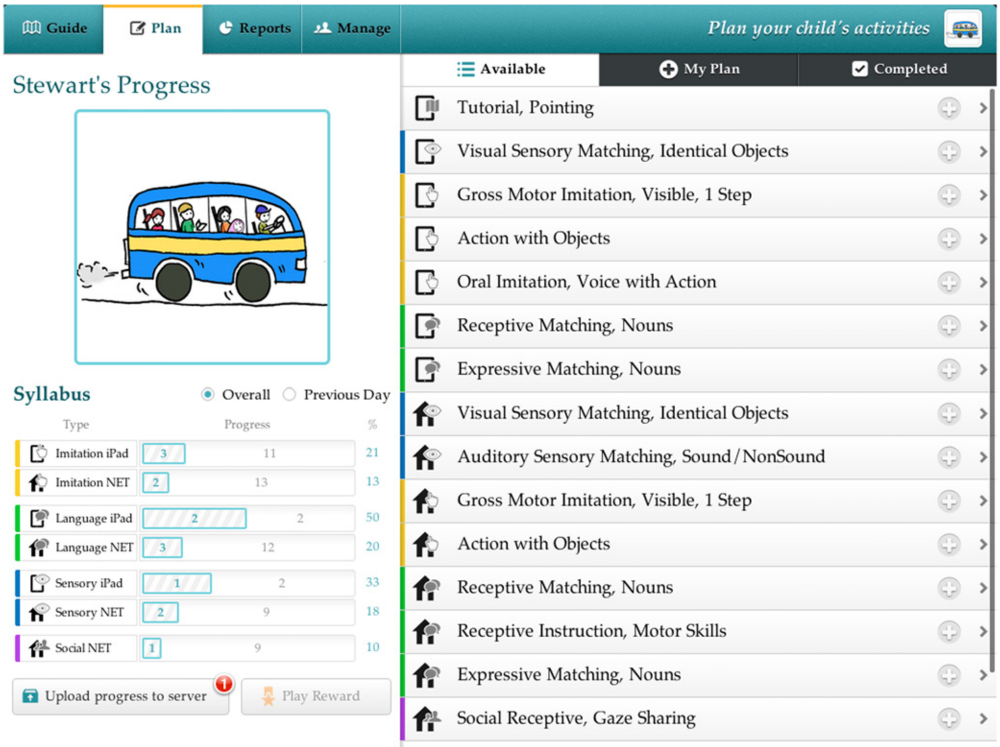


Fig. 4. Screenshot of Playpad's daily schedule. Left of screen includes details of the current child's profile and progress; Right of screen lists tasks that Playpad has made available.

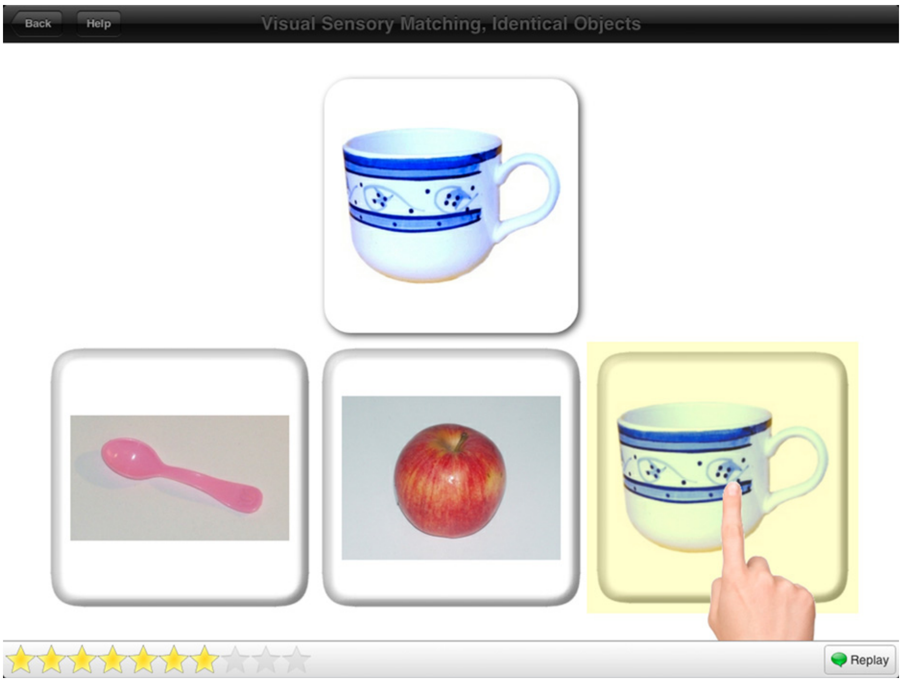


Fig. 5. Screenshot of visual sensory matching. The child must match pictures using only sensory cues. Here we show the system prompting the correct response.

Symbolic processing — distinction based on categories: or "What does not belong?" The system presents three exemplars from one semantic category, and one example from a different category, and the child must discern the odd one out. The

Fig. 6. Screenshot of symbolic processing task. The child must identify which object (penguin) is different from the others (kangaroos).

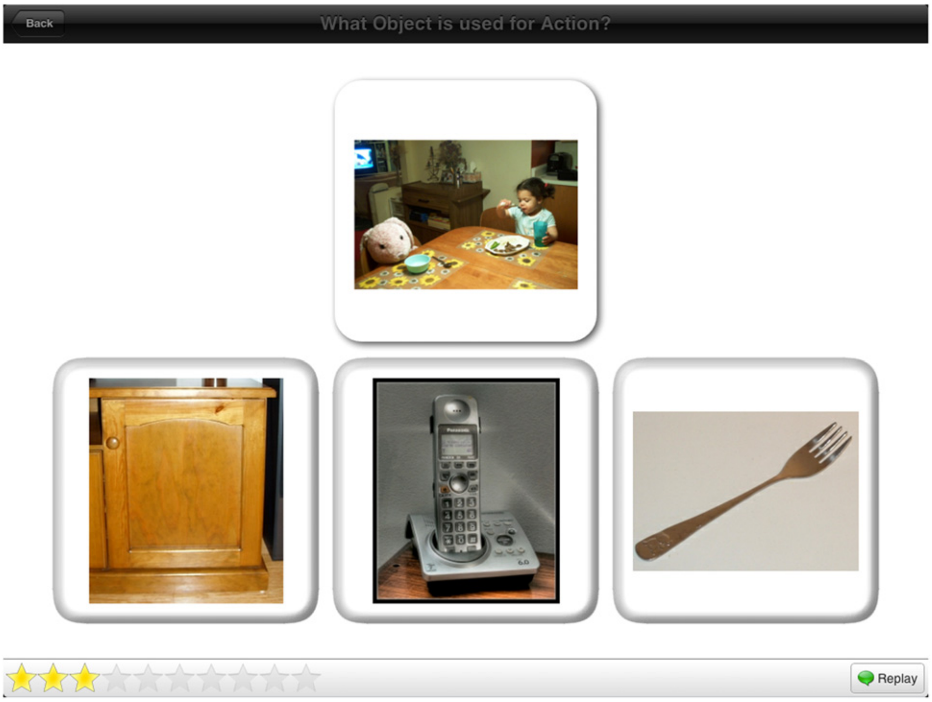


Fig. 7. Screenshot of functional matching task. The subject must identify which object (fork) is used for the action (eating).

system uses closeness in the semantic network as a measure of similarity. Thus "penguin" and "kangaroo" are close because they are both sub-types of "animal." Categories that are semantically close are deemed more "natural" comparisons, but also more difficult to discriminate. An example is shown in Fig. 6.

Functional matching: the child is presented with an action, and must choose the target (an object used to perform that action) from two distractors (objects not used to perform that action). To do this, the system leverages information in the semantic network regarding functional uses of objects. For example "a broom is used for sweeping", "a pencil is used for writing". The task can be generalised over any semantic relation represented in the network (e.g., "found in location", "can perform action", "is part of"). An example is shown in Fig. 7.

6.2. Receptive and expressive language tasks

Screenshots of matching tasks are shown in Fig. 8. The left screen is a session of visual receptive matching. The stimulus is the utterance "find apple". The child must identify the item that matches "apple". The right screen shows an expressive matching session. The stimulus category is apple, and the instruction given is "What is this?". The child must say "apple", and the parent uses the feedback buttons to indicate the response ("Yes", "Prompt", "No").

6.3. Natural environment tasks

A screenshot of a NET task is shown in Fig. 9. For this task, the child must find items from a shopping list. At the top left is a scrollable list containing instructions to the parent, and below the list are the items which can be checked off according

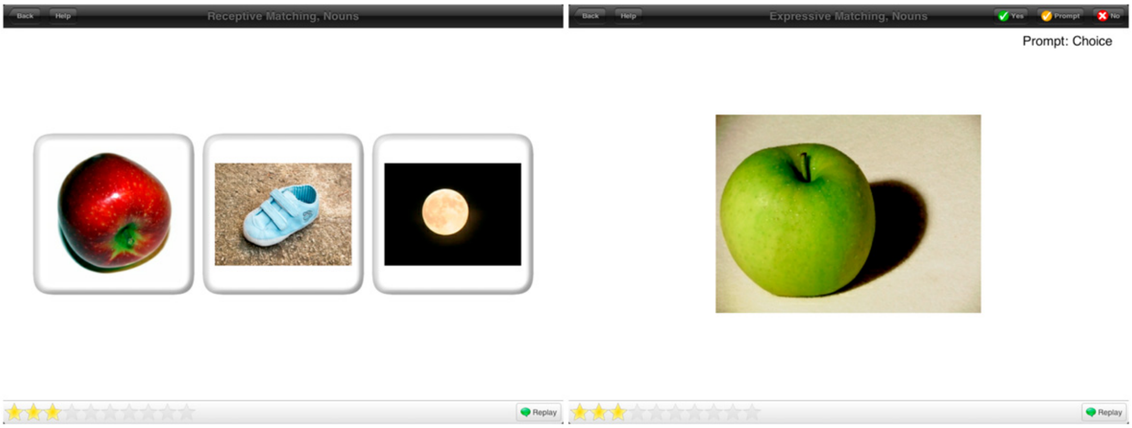


Fig. 8. Screenshots of visual receptive matching (left) and expressive matching (right).

Fig. 9. Screenshot of natural environment task "Let's Go Shopping".

to whether the child is able to find them or not. At the top right are instructional videos demonstrating how to perform the task.

NET activities are intended to be done in the real world, with Playpad playing a supporting role, providing resources to facilitate learning of the task, and collecting feedback from the parent. Consider two examples from our system:

Real world activity, responses on iPad — What's in the bag?

1. Sit on the floor or table with your child. 2. Take out one item at a time, holding it up to show your child. Encourage them to label the food item, e.g., "What's that?", "Look!", "It's a ______", point to the item and look expectantly at your child (use your prompt techniques). 3. Set the food item aside once it has been labelled; to extend your child's skills, group the

Table 1
Stimuli evaluation by therapists: Meanings of ratings for the first item, *Stimulus array levels of complexity–Colour*, are: 1–Did not adapt; 2–Some evidence of adaptation but far from ideal; 3–Somewhat adapts; 4–Partial adaptation but not yet ideal; and 5–Ideal and exact adaptation. Ratings for the other items were interpreted similarly.

Aspect rated (scale 1–5)	Category	T1	T2
Stimulus array levels of complexity	Colour	5	5
Stimulus array levels of complexity	Shape	5	5
Stimulus array levels of complexity	Object	5	3.5
Sequences of stimulus presentation demand	After error	4	4
Sequences of stimulus presentation demand	Controlling for position	5	5
Sequences of stimulus presentation demand	After mastery	4	4
Rate prompt	Accuracy	5	5
Rate prompt	Clarity	3	3.5
Rate reinforcement	Clarity	4	4
Rate reinforcement	Continuity	4	5

items according to food group (fruit in one pile, vegetables in another pile). 4. You may like to offer your child a turn at taking something out of the bag. 5. Continue until all items have been taken out of the bag. 6. Record the required prompting level once finished.

Mixed real-cyber activity, responses on iPad — Let's go shopping.

1. Select the "Let's Go Shopping" game on Playpad. 2. Take your shopping bag and Playpad into the kitchen. 3. Select one item from the "Shopping List" displayed on Playpad, listening to the name of the item. 4. Select the picture on Playpad again to hear the name of the item to help match the spoken word with the object and/or the picture. 5. Encourage your child to independently take the item and place it in the shopping bag. 6. Record on Playpad whether your child independently matched the real object with the picture on Playpad (select "yes", "no", "prompt"). 7. Playpad will take you back to the "Shopping List" page. Continue until you have collected the foods on the list. 8. Take the food you have collected and have your meal.

7. Experiments and evaluation

This section contains the results of our evaluation, both qualitative and quantitative, from all three stakeholders–parents, therapists and children.

7.1. Qualitative assessment of automated stimulus generation

The quality of the stimuli automatically generated for matching tasks was evaluated by two therapists (T1 and T2) having more than 15 years experience in therapy. In particular, they examined the visual-to-visual matching, "What does not Belong", and "Functional matching" tasks. Their evaluation is found in Table 1.

The evaluation measures different aspects of stimulus generation, prompting, and reinforcement. *Stimulus array levels of complexity* assesses if the complexity of stimuli increased accurately, and is tested for three categories: colour, shape and object identity matching. *Sequences of stimulus presentation demand* assesses if stimuli are correctly readjusted under three conditions: after error, after mastery, and after controlling for position. *Rate prompt* assesses the accuracy and clarity of the supplied prompting. *Rate reinforcement* assesses the clarity and continuity of the supplied reinforcement. Therapists ratings are given on a scale from 1 (worst) to 5 (best). The meanings of ratings for the first item in Table 1, *Stimulus array levels of complexity–Colour*, are: 1 – Did not adapt; 2 – Some evidence of adaptation but far from ideal; 3 – Somewhat adapts; 4 – Partial adaptation but not yet ideal; and 5 – Ideal and exact adaptation. Ratings for the other items were interpreted similarly.

7.2. TRIAL 1: Evaluation of the instructional design with autistic children

Seven children diagnosed with Autism (6 boys and 1 girl; aged 4 to 8; and one 14, mentally aged 7) participated in the first formal study involving the intended users, together with a 3 year old with no autism. Participants performed tasks on the iPad for 30–60 min. This experiment investigates if mastery is being achieved, if prompting is adjusted to stimuli, and if retraining occurs correctly.

The design of the experiments is as follows. Two categories of unseen symbols are created ($a0$, $a1$, $a2$) and ($b0$, $b1$, $b2$). These symbols can represent any stimulus drawn from the categories used in this paper, and are equivalence classes formed from arbitrary collections of objects, or from rules, e.g., same colour, same function.

Training involves repeated presentations of the following two sets: $\{(a0, a1)(b0, b1)\}$, $\{(a1, a2)(b1, b2)\}$. If the subject makes a mistake during training, they receive feedback and the correct response is indicated. The testing phase examines whether the concept equivalence has been learned and can be transferred in a transitive manner as $\{(a0, a2); (b0, b2)\}$. These derived relations are not presented during training, and the user is not given feedback during testing that would

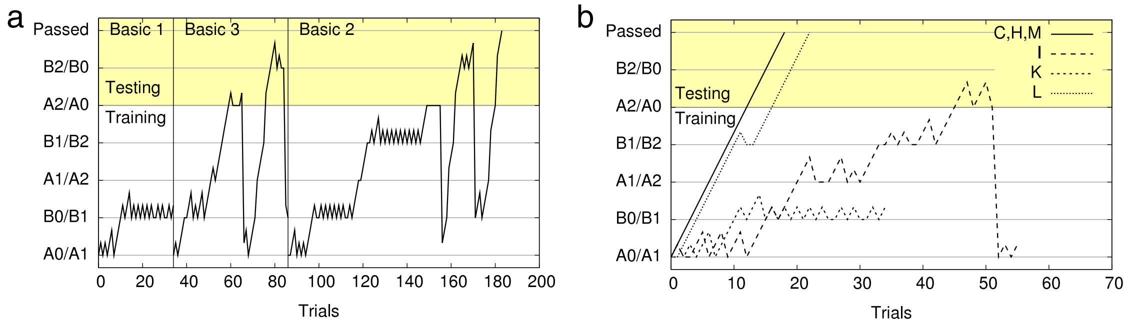


Fig. 10. (a) Results for autistic user "K", who initially failed to complete the tasks but improved with practice, and was able to finish the final task; (b) Results for 6 autistic users on the "Basic 1" task. Some users completed the task without errors or prompting (C, H, I). Others (I, K) were unable to finish the task.

otherwise reinforce these relations. If the derived relations are not correctly matched, the system re-presents the base relations.

Fig. 10(a) plots results for an autistic user K on the simple tasks of choosing one of two shapes based on colour or size. The shaded, upper part of the graph represents testing, and the lower part represents training. The *y*-axis represents the number of consecutive correct answers for a stimulus pair, the *x*-axis represents number of trials. A user is considered to have mastered a pair after five consecutive correct answers, whereupon training moves to the next pair. Consider the cyan (rightmost) series: the user is presented with A0/A1, B0/B1, A1/A2, and B1/B2, before moving into testing. The user gets the derived relation A2/A0 wrong, so the system moves into retraining (A0/A1 followed by A1/A2). After then passing A2/A0, the user fails at B2/B0, and the system retrains with B0/B1, followed by B1/B2. It then retests B2/B0, which the user has now mastered.

The tasks user K found difficult were performed by three other users with no errors or prompting, but there was a wide spread of ability across autistic users. E.g., Fig. 10(b) shows results for 6 users on the simple task of choosing one of two shapes based on colour. Users C, H, and M performed this task with no errors or prompting. Users I and K were unable to finish the task, however user K, after attempting three such tasks, went on to complete the final task by choosing between two dots based on size.

The overall results were very positive, and are consistent with the outcomes and the adaptive learning processes established in studies of stimulus equivalence [18]. As such the protocol, presentation and medium utilised in this study can be considered a successful replication of empirically supported training protocols. The system was able to identify how different learners struggled at certain sub-components and target those components by returning to training levels (see Section 7.5 for measures indicative of learning). Moreover, the detailed data collected by Playpad allowed instructional designers new insights, enabling them to subsequently refine the methods. For example, the system was able to identify that one child discerned equivalence on some stimuli but not others. As a result of this information, that the child's therapists are now able to adjust her program based on the revealed sensory bias, and the consequences this will have for future cognitive learning.

7.3. TRIAL 2: Evaluation of matching, expressive language and natural environment tasks

A detailed follow-up trial was performed with eight children with a diagnosis of Autism (7 boys and 1 girl; ages 2 to 7) and their parents over two weeks. Each child performed an hour long session on the iPad, including 20 min on-iPad activity and 20 min of NET every day in that 2 week period. Two sets of tasks were targeted: matching for 4 children with early diagnosis, and expressive tasks for 4 children who have passed the receptive language stage. In total the children performed about 250 activities over a two week period, with around 2900 discrete trial responses being recorded in the system. Below we detail the results.

7.3.1. Adaptation of the system during individual tasks

The trial includes two types of matching tasks. The system presents an array of three pictures, only one of which is correct. For receptive matching, the child is instructed via a verbal or visual clue, and indicates their response by touching an image from the array of choices. The computer controls stimulus generation, recording of response, prompting, reinforcement and (if necessary) rote training. For expressive matching, the child is presented with a single picture and is asked the question "what is this?". They respond verbally by saying the name of the object, and the parent or therapist evaluates and responds with "Yes", "No", or "Prompt". In both cases, the computer adapts the selection of stimuli and reinforcement to the performance of the subject.

The relationship between prompting, reinforcement, and errors is illustrated in Fig. 11. This shows performance in one session of the "Colours" task, in which the subject must select one of three images that matches the colour of the stimulus image. The task includes several stages of increasing complexity, initially with no distracting factors, then later introducing

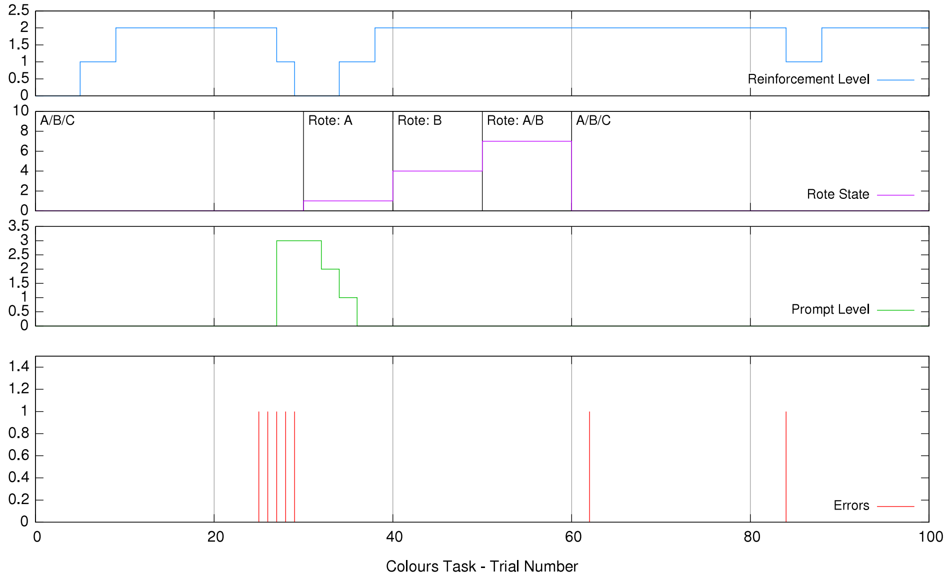


Fig. 11. Subject performance for *colours* task showing rote cycle.

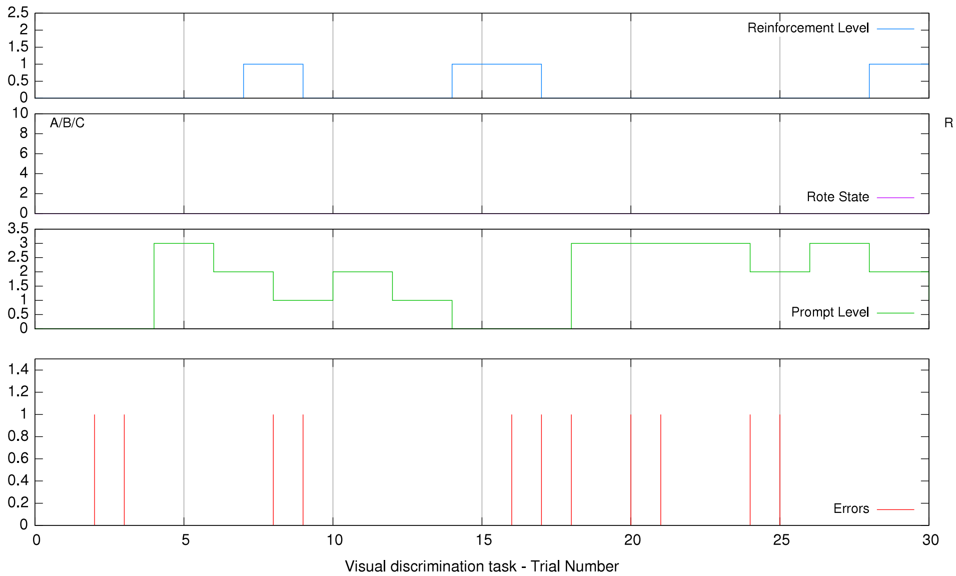


Fig. 12. Subject performance for "visual sensory matching" task. High error rates cause high levels of prompting and reinforcement.

varying abstract shapes, and finally coloured objects. The top three timelines show reinforcement level, rote training state, and prompt level. Errors are indicated as impulses on the bottom chart. Reinforcement is delivered in the form of animated fireworks, initially with every correct response (level 0). After sufficient correct responses, reinforcement fades to every two correct responses (level 1) and then randomly every 1 to 5 responses (level 3). Reinforcement increases in the presence of errors (e.g., at trials 27, 29 and 84). Prompting is introduced after two consecutive errors. It begins at maximum level (level 3), highlighting the correct response by pulsing, circling and pointing. Prompting fades with correct responses, and increases with incorrect responses, adapting to the needs of the subject. After three consecutive errors at maximum prompt level, the system enters rote training (at trial 30). This involves repeated presentations of the original stimulus (A), then an alternative stimulus (B), then a mixture of the two (A/B), before returning to the original A/B/C stimulus schedule. These states are indicated on the chart.

The previous example shows relatively low levels of reinforcement and prompting. Fig. 12 shows how more frequent errors result in greater prompting and reinforcement. This is for the "Visual sensory matching" task, in which the subject must identify which of an array of three images duplicates the stimulus image. This task requires no understanding of categories, simply the ability to visually discriminate identical images. The task begins with coloured abstract shapes

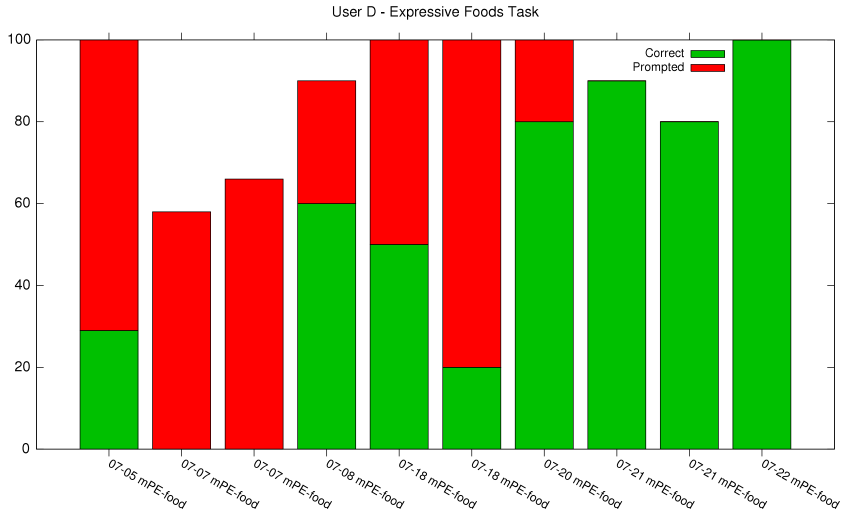


Fig. 13. Improvement in performance over time of subject D for "expressive matching food" task. Duplicate labels indicate a task was attempted multiple times on the same day (e.g., 07–07 mPE-food).

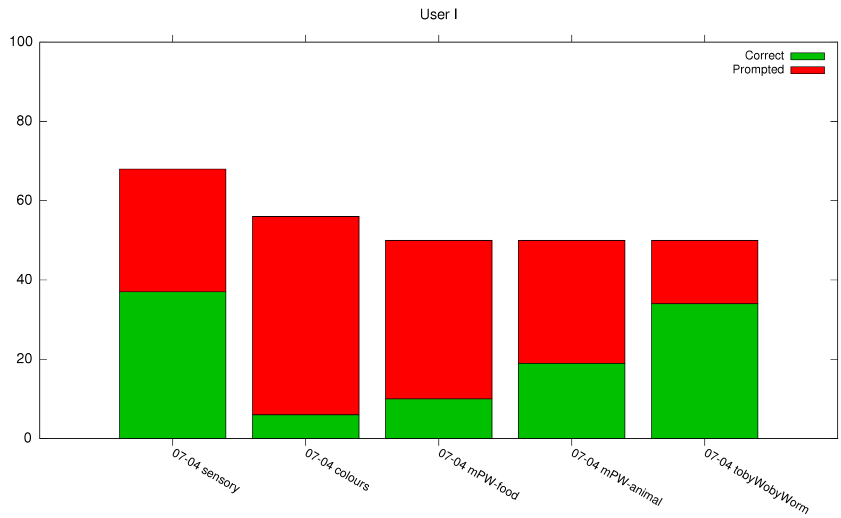


Fig. 14. Improvement in performance over time for subject I.

(e.g., crescents, triangles), and progresses to animals and household objects. The subject answered 64% of the trials successfully. Frequent errors caused a high level of prompting — 74% of all responses and 51% of the correct responses are prompted.

7.3.2. *Examples of individual performance*

We now detail examples of individual performance. With reference to Figs. 13–16, the *x*-axis indicates date and name of task for each trial. The name of task is split as Sensory matching (Colours), mPW (Receptive matching — picture to word), mPE (Expressive matching), and NET (Picnic time, Yummy in my tummy, What's in the bag, Let's go shopping, Toby Woby worm, Birthday party). The height of each bar in a chart indicates the proportion of responses that were correct. The green portion indicates unprompted responses and the red portion prompted responses.

Improved performance: As learning occurs we expect to see two things. First, a decreasing number of errors, and increasing proportion of correct responses. Second, a decreasing level of prompting. Fig. 13 charts an example, this time of the expressive task "matching food". During this task, the subject is shown pictures of common foods (e.g., apple, banana, carrot, cheese), and is asked the question "What is this?". The parent checks the child's response, entering "Yes" (correct), "No" (incorrect) or "Prompt" (correct with prompt). Otherwise, the system responds in the same way as for receptive trials, by varying the level of reinforcement, and adjusting the stimulus as required. Over time the level of unprompted responses rises. (See also Fig. 14, where although performance is constant, the level of prompting decreases). We aim for "errorless

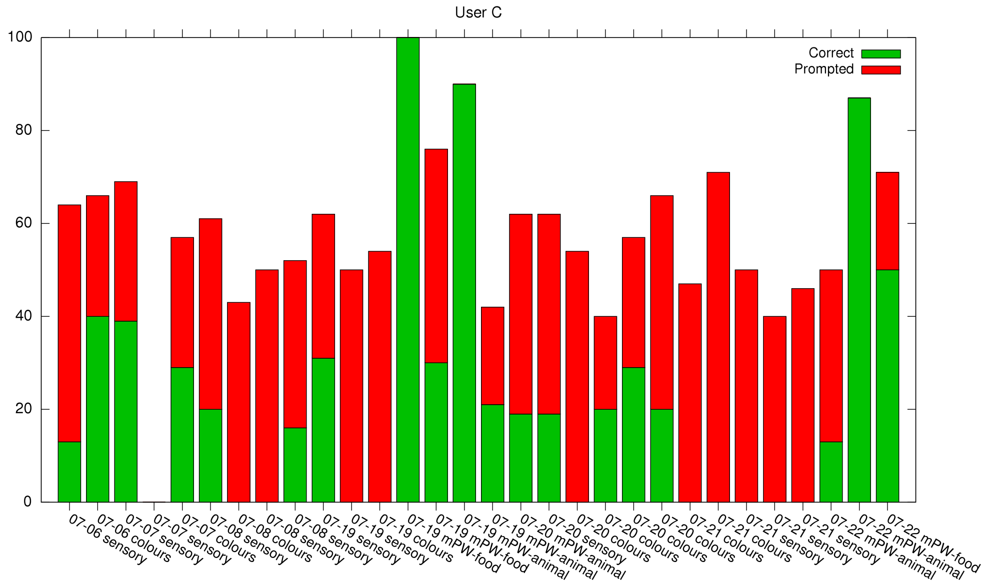


Fig. 15. Results for subject C, showing difficulty with simpler receptive matching tasks.

Table 2
Exit survey: Parent teaching experience.

Parent teaching experience — scale (1–5)	A	B	C	D	E	F	G
Did you enjoy using the iPad?	5	5	5	4	5	5	4
Were the teaching activities effective?	3	5	5	4	5	5	4
Did you feel empowered in teaching?	3	5	5	4	5	5	4
Did you feel you learnt the target activities over the week?	Partially	Yes	Yes	Yes	Yes	Yes, easy	Partially
Did you start to think of your own NET activities?	NA	Yes	NA	No	No	No	NA

Table 3
Exit Survey: Parent evaluation of child learning experience (Parent G: Child distracted in new environment).

Parent evaluation of child learning — scale (1–5)	A	B	C	D	E	F	G
Did you feel your child enjoyed the iPad?	2	5	5	4	5	5	3
How easy was the iPad to use?	5	5	4	5	5	5	3
How much use would the iPad be for teaching your child?	3	5	5	4	5	5	5
How much use would the iPad be for entertaining your child?	3	5	5	4	5	5	4
Did you feel your child learned the skill over the week?	No	Yes	Partially	Yes	Yes	Yes	No

learning'', a form of conditioning that minimises negative stimuli (e.g., absence of reinforcement). Prompting allows the child to receive higher reinforcement during training, increasing their engagement with the process.

Struggling children: This trial included children of a range of abilities. While some were able to eventually master expressive matching tasks, others struggled with even basic receptive and sensory matching. Fig. 15 charts the performance for a user across a number of receptive matching tasks, including the ''Visual sensory matching'' and ''Colours'' tasks already described, as well as receptive matching of common foods and animals. In these tasks the subject is shown a set of pictures, and is verbally instructed to find one of them (e.g., ''Find cat'', or ''Find banana''). The response is indicated by touching the corresponding picture. While performance is occasionally good for these tasks, the majority of responses required much prompting, indicating that the skills have not been mastered.

Excellent performance: Some users displayed excellent performance, even on complex expressive tasks. Examples of their results are shown in Fig. 16.

7.4. Exit survey from parents

Parents were asked to evaluate both the quality of their teaching experience with Playpad and their child's experience in learning. Results of this survey are found in Tables 2 and 3. The responses are overwhelmingly positive in all aspects being assessed, both for the child and the parent. Of the seven parents who returned the survey, only one (A) reported an average experience; all others rated most experiences with the highest score.

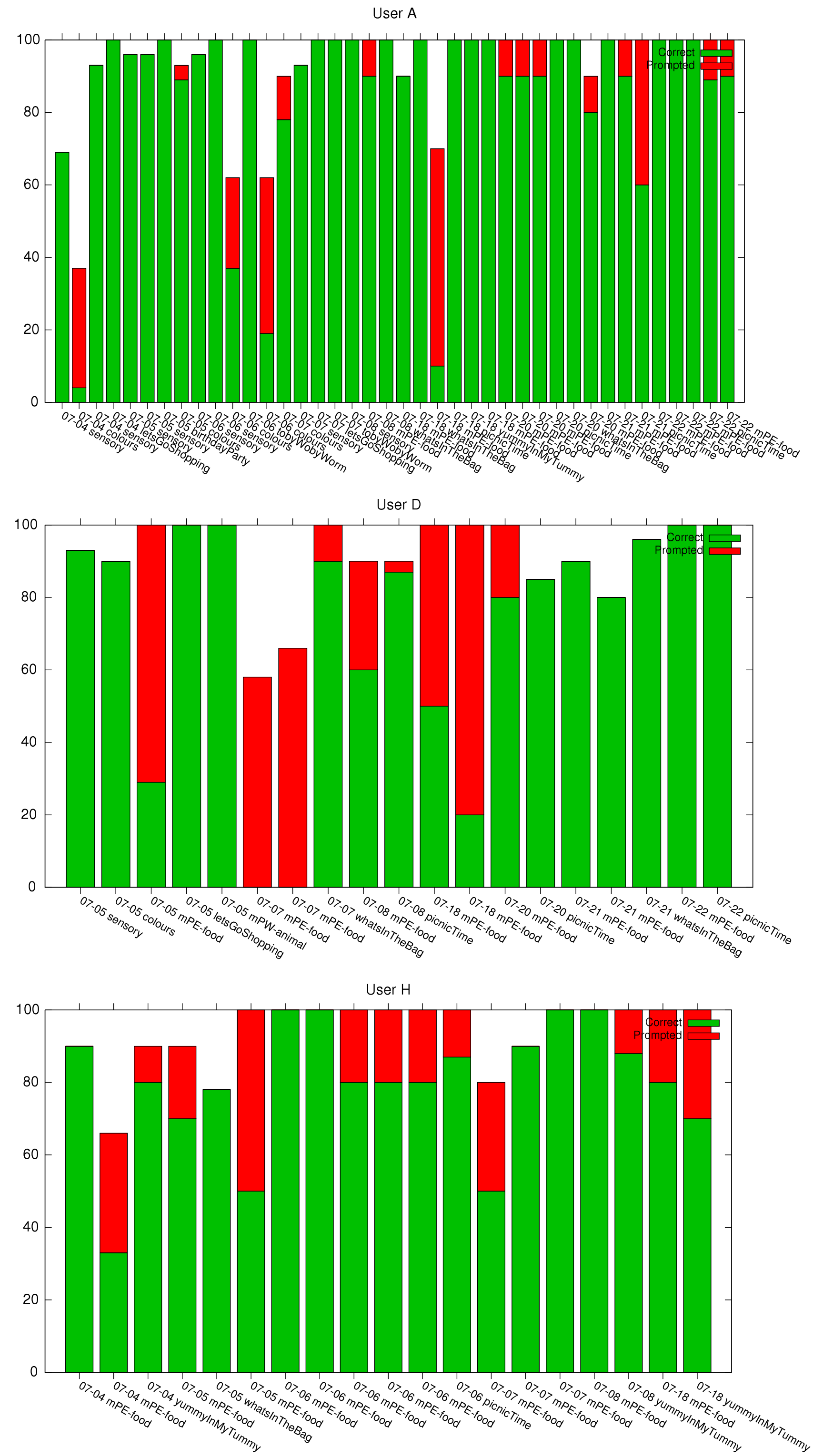


Fig. 16. Results for subjects A, D, and H, showing good performance on a variety expressive matching and NET tasks.

7.5. TRIAL 3: Longitudinal field trial

Encouraged by the results obtained in the initial trials, we undertook a third, larger trial under conditions that match the envisaged conditions of use of Playpad. This trial consisted of 16 participants (12 boys and 4 girls) and their parents,

Table 4
Trial 3 learning progress for all tasks across 16 children.

Start state to end state	Number of tasks
Fail to Fail*	21
Fail to Prompt	6
Fail to Success	18
Prompt to Prompt (same level)*	22
Prompt to Prompt (lower level)	38
Prompt to Success	98

* Only the criteria *Fail to Fail* and *Prompt to Prompt (same level)* demonstrate no learning; all other criteria indicate some learning has occurred.

not part of the previous trials, and ranging in age from 3 to 15. The trial extended over a 4 week period. Participants had no formal introduction to the software and navigated through the syllabus at their own pace. Below we present a summary of the participants' progress over this period with an emphasis on evidence of learning. Learning is measured according to the following 6 criteria:

(i) Fail to Fail: where the child fails the first attempt at a task, and fails the last attempt at the same task;
(ii) Fail to Prompt: where the child fails the first attempt at a task, but completes the same task with prompting on the last attempt;
(iii) Fail to Success: where the child fails the first attempt at a task, but completes the same task without prompting on the last attempt;
(iv) Prompt to Prompt (at same level): where the child completes the first attempt at a task with prompting, and the last attempt at the same task with the same level of prompting;
(v) Prompt to Prompt (at lower level): where the child completes the first attempt at a task with prompting, and the last attempt at the same task with a lower (i.e., better) level of prompting;
(vi) Prompt to Success: where the child completes the first attempt at a task with prompting, and finishes the same task at last attempt without prompting.

Table 4 shows aggregates for these criteria across all tasks and children. Criteria (i) and (iv) indicate no learning, whilst all other categories indicate learning. Participants were able to learn from the first failed attempt at a task. On average, across all participants, 13% of tasks initially failed $(6/(21+6+18))$ were eventually completed with some level of prompting, whilst 40% of tasks intially failed were completed without any prompting. It is encouraging to see that the majority of tasks intially requiring prompting were either ultimately completed without prompting: $98/(22+38+98) = 62\%$, or else completed with a *lower* level of prompting. Playpad tracks this progress at a level that is, to the best of our knowledge, unprecedented.

7.6. Discussion

Our results demonstrate categorically that the instructional framework is suited to learning, in terms of how stimulus is adjusted for retraining and how performance improves over trials and across concepts.

The impact of the teaching system's instructional design, task sequencing, and stimulus generation can be seen in its ability to deliver complex teaching tasks that adapt to the needs of learners while remaining easily adapted to therapists' goals. The quality of the automated stimulus complexity generated was evaluated qualitatively by experienced therapists with very positive feedback. Importantly, this sentiment was echoed from the field by those who are the primary users of the system: the parents were overwhelmingly positive about the their teaching experience and their children's experience of using this learning environment. In short, to quote one child: "Mummy! The teacher is in this [iPad]!"

The most touching result was achieved in week 2 of Trial 2, when one child asked a question for the first time: "What is this?" We consider this the highlight of the trials, because teaching an autistic child to ask a question — something fundamental to the process of learning — is extremely difficult. We note that Playpad was not explicitly teaching this skill; it was acquired as a result of the learning environment.

Areas identified for improvement are the delivery of NET tasks. The difficulties encountered were: (i) The NET tasks were not explained to the parent, and thus they did not quite understand the role of the task; (ii) The prompting protocol in NET tasks needs to be elaborated to the parent; and (iii) The recording on Playpad was challenging in NET tasks, as the parent struggled to control the teaching environment.

8. Conclusion

We have described and evaluated a framework for pervasive delivery of early intervention therapy to children with ASD. The system is able to deliver stimulus according to therapists' goals and adapt automatically to a child's performance. Evaluation with adults and target users has demonstrated how the system is able to accurately adjust to individual performance, and how all subjects transfer learning across stimulus sets. The accuracy of stimulus generation and response

recording opens new avenues in learning and early intervention. The best comment received was from a 4 year old who participated in our study: *Mummy! The teacher is in this [iPad]*! Ongoing trials are confirming the value of pervasively-delivered therapy that can be carried at all times, and handed between parents, therapists, and teachers. Surprisingly, we are finding that for some children, and some competencies, the Natural Environment Tasks are aiding on-iPad tasks. Future work will target extension of the system with new category schemata, under the direction of our partner ASD therapists, and larger, longitudinal trials of Playpad in its intended role as a wait-list technology.

References

[1] Department of Health. Center for disease control and prevention, autism and developmental disabilities monitoring (ADDM) networks (Atlanta, USA). Online: http://www.cdc.gov/ncbddd/autism/about.html.
[2] Mark O. Riedl, Rosa Arriaga, Fatima Boujarwah, Hwajung Hong, Jackie Isbell, L. Juane Heflin, Toward assisted authoring of social skill scenarios for young adults with high functioning autism, in: Proceedings of the IJCAI Workshop on Assisted Cognition, California, 2009.
[3] J.A. Kientz, G.R. Hayes, T.L. Westeyn, T. Starner, G.D. Abowd, Pervasive computing and autism: assisting caregivers of children with special needs, IEEE Pervasive Computing (2007) 28–35.
[4] K. Blocher, R.W. Picard, Affective social quest: emotion recognition therapy for autistic children, in: In Socially Intelligent Agents-Creating Relationships with Computers and Robots, Kluwer Academic Pub., 2002.
[5] G. Montes, J.S. Halterman, Association of childhood autism spectrum disorders and loss of family income, Pediatrics 121 (4) (2008) e821.
[6] C. Schmidt, M. Schmidt, Three-dimensional virtual learning environments for mediating social skills acquisition among individuals with autism spectrum disorders, in: Procs. of the Int. Conf. on Interaction Design and Children, ACM, 2008, pp. 85–88.
[7] C.M. Stanton, P.H. Kahn Jr, R.L. Severson, J.H. Ruckert, B.T. Gill, Robotic animals might aid in the social development of children with autism, in: Procs. of the ACM/IEEE Int. Conf. on Human Robot Interaction, ACM, 2008, pp. 271–278.
[8] V. Bernard-Opitz, N. Sriram, S. Nakhoda-Sapuan, Enhancing social problem solving in children with autism and normal children through computer-assisted instruction, Journal of Autism and Developmental Disorders 31 (4) (2001) 377–384.
[9] L. Schreibman, C. Whalen, A.C. Stahmer, The use of video priming to reduce disruptive transition behavior in children with autism, Journal of Positive Behavior Interventions (2000).
[10] A. Bondy, L. Frost, A. Bondy, A Picture's Worth: PECS and Other Visual Communication Strategies in Autism, Woodbine House, 2001.
[11] P.A. Lorimer, R.L. Simpson, B. Smith Myles, J.B. Ganz, The use of social stories as a preventative behavioral intervention in a home setting with a child with autism, Journal of Positive Behavior Interventions 4 (1) (2002) 53.
[12] A. Bosseler, D.W. Massaro, Development and evaluation of a computer-animated tutor for vocabulary and language learning in children with autism, Journal of Autism and Developmental Disorders 33 (6) (2003) 653–672.
[13] O.E. Hetzroni, J. Tannous, Effects of a computer-based intervention program on the communicative functions of children with autism, Journal of Autism and Developmental Disorders 34 (2) (2004) 95–113.
[14] O.I. Lovaas, Behavioral treatment and normal educational and intellectual functioning in young autistic children, Journal of Consulting and Clinical Psychology 55 (1) (1987) 3–9.
[15] M.M. Mueller, C.M. Palkovic, C.S. Maynard, Errorless learning: review and practical application for teaching children with pervasive developmental disorders, Psychology in the Schools 44 (7) (2007) 691–700.
[16] S. Berkowitz, A comparison of two methods of prompting in training discrimination of communication book pictures by autistic students, Journal of Autism and Developmental Disorders 20 (2) (1990) 255–262.
[17] M.E. Libby, J.S. Weiss, S. Bancroft, W.H. Ahearn, A comparison of most-to-least and least-to-most prompting on the acquisition of solitary play skills, Behavior Analysis in Practice 1 (1) (2008) 37.
[18] R.R. Saunders, G. Green, A discrimination analysis of training-structure effects on stimulus equivalence outcomes, Journal of the Experimental Analysis of Behavior 72 (1) (1999) 117.