MASTER'S THESIS
Learning in complex systems
February 2018

Tablet Technology in Applied Behavior Analysis

Vetle Berge

Faculty of Health Sciences
Department of behavioral science

OSLO AND AKERSHUS
UNIVERSITY COLLEGE
OF APPLIED SCIENCES

Acknowledgements

I would like to express my sincerest appreciation to Sigmund Eldevik, PhD for excellent guidance and supervision throughout the process of my master thesis. I would also like to thank professor Svein Eikeseth, PhD, Børge Strømgren, PhD, and the rest of the applied and experimental behavior analysis in clinical practice research group for continuous discussion and feedback. Lastly I would like to give special recognition to Marie Moksness and the rest of Superplus for their cooperation, as well as thank my family and friends for their continued support.

Table of contents

Article 2

Establishment and Generalization of Receptive Labeling following Tablet-based Instruction in Children with Autism

List of tables and figures

Article 1

Article 2

A Quantitative Review of Tablet-based Technology in Applied Behavior Analysis

Vetle Berge

Oslo Metropolitan University

Abstract

Recently there have been published several studies employing tablet-based interventions in the treatment of children with disabilities. Tablet-based interventions have several advantages over traditional methods, especially because they are mobile, customizable, and are easy and intuitive to maneuver. However, due to the novelty of these interventions there seems to be a lack of studies evaluating the effectiveness of tablet-based interventions. As a result of this, the current study set out to conduct a quantitative review of the effectiveness of studies using tablet-based interventions. This was accomplished by applying the recently developed single subject research effect size, non-overlap of all pairs. The results indicated that, on average, tablet-based interventions had a medium effect. However there was a large verity between studies, indicating that tablet-based interventions might not be a one-fits-all solution and that future research should further evaluated which factors that might affect the effectiveness.

Keywords: Tablet, technology, Non-overlap of All Pairs, Autism Spectrum Disorder, Meta-analysis, quantitative review

A Quantitative Review of Tablet-Based Technology in Applied Behavior Analysis

Technology is developing quickly and has rapidly become an essential component of our daily lives. While the computer made an entry into our workplaces and homes in the nineties, the tablet computers are now swiftly playing a similar role in both our social lives and work. Due to these recent technological advances computers have in the last decades been utilized in the treatment of children with disabilities. Several studies have investigated the use of computer-based interventions, and have found that computers have been utilized in teaching several skills; academic (Basil & Reyes, 2003; Hetzroni, Rubin, & Konkol, 2002; Knight, McKissick, & Saunders, 2013; Yamamoto & Miya, 1999), social (Ayres & Langone, 2002; Simpson, Langone, & Ayres, 2004), and matching (Shimizu, Twyman, & Yamamoto, 2003). A report completed by a Norwegian government agency in 2016 found that tablet computers are now as widespread for today's youth as traditional computers. Their report indicated that 96,1 % of 7^{th} grade students have access to a computer at home with 64,1 % of them being in possession of their own. 90,7% of the same students reported to have access to a tablet, with 68,5% owning one (Egeberg, Hultin, & Berge, 2017). As a result of this large-scale adaptation, tablet computers are currently entering the field of behavior analysis, especially with regard to the treatment of young children with developmental disabilities like autism spectrum disorder (ASD).

Early intensive behavioral intervention (EIBI) are by many viewed as the most effective treatment for young children with ASD and as been proven to significantly improve the rate of learning (Klintwall et al. 2015) and intellectual and adaptive functioning (Eldevik et al., 2010; Virués-Ortega, 2010). EIBI is introduced between two and six years of age. The intervention targets multiple skills of clinical significance, which usually includes skills such as eye contact, imitation, non- vocal and vocal communication, self-care, social interactions, pre academics, fine and gross motor skills, play and leisure (Green, 2011). The intervention is highly individualized; in the sense that the objectives are tailored to the child's strengths and weaknesses. EIBI training is usually conducted in multiple environments; generally starting in the home of the child before gradually moving to a school setting (Green, 2011).

Tablet computers have several advantages over the traditional laptop or desktop computer. Tablet computers are mobile, meaning that the training does not have to occur in a clinic setting (Allen, Hartley, & Cain, 2016). This is significant when it comes to the treatment of children with ASD due to the fact that training is recommended to occur in multiple settings, and with high intensity (Green, 2011). Tablet computers are also often more customizable and flexible then the traditional computer, which suits EIBI due to the necessity of highly individualized interventions and the targeting of several skill domains (Green, 2011). Another advantage of tablet computers is that they are usually less expensive compared to traditional computers, thereby making them more accessible. Tablet computers are also generally viewed as more suited for young children because they are considered more intuitive and easy to navigate compared to traditional computers. Another considerable advantage of implementing tablet computers in applied behavior analysis is that the data can be stored electronically and automatically, thereby optimizing the precision of the data recording as well as allowing effective time management (Artoni et al., 2013).

Recent studies comparing the application of tablet computers with traditional training have indicated that tablet computers can promote the rate of improvement (Allen et al., 2016) as well as decrease the amount of task refusal (El Zein et al., 2016). Knight et al. (2013) also point out that children might prefer technology-based interventions, as it utilizes a platform usually associated with leisure and resemble a game. Technology based interventions might also have the advantage of facilitating autonomy and allow teachers to work with multiple children at a time and thereby potentially be more time-effective then traditional interventions (Knight et al., 2013). Another important consideration in regards to implementing technology for children with disabilities is that today's children need to be thought to use of technology to be successful workers in the future (Knight et al., 2013).

The conceivable advantages of using tablet computers in training with individuals with ASD are therefore compelling. However, there is a need to evaluate the effectiveness of the applications available. Tincani and Boutot (2005) emphasized for instance the necessity of determining the effectiveness of technology for children with ASD before the implementation in training programs. Computer based interventions have already been thoroughly reviewed (Knight et al., 2013), however there seems to be a lack of investigations evaluating tablet technology especially when it comes to thoroughly assessing available applications. This might partly be caused by the controversies regarding the use of single case research in quantitative reviews.

There are several methods available to assess the effectiveness of studies employing single subject research designs. Parker, Hagan-Burke, and Vannest (2007) sampled 75 studies deploying single subject research designs, and found that 87% relied solely on visual analysis, thereby supporting the notion that visual inspection is the main tool for evaluating the effect of interventions in behavior analysis. Visual inspection consists of evaluating the consistency of the intervention effect by visually assessing graphical data (Kazdin, 2011). The visual analysis is often assisted by comparisons of phase means, medians or percentages, but rarely with any statistical tests of differences between these indices (Parker et al., 2007). This method is ordinarily less sensitive towards detecting small intervention effects, which is not necessarily considered an limitation in behavior analysis as it leads to the uncovering of only the most powerful and consistent intervention effects (Kazdin, 2011). In fact Baer (1977) argued that the application of visual inspection is a fundamental component of behavior analysis, as it contributes to a more concise technology of behavior and concluded that it should not be abandoned. Parker, Cryer, and Byrns (2006) supplement that point of view by arguing how no statistical technique can match the holistic and integrative nature of visual analysis when it comes to considering data variability, trend magnitude and direction, mean levels and shift, and embedded cycles in performance. They also emphasize how effect sizes only are an index of change, hence not able to distinguish between improvement and deterioration, a consequential limitation when it comes to evaluating the effect of interventions on socially significant behavior, where the direction of change is especially essential (Parker et al., 2006).

However there are some instances where the quantification and statistical analysis of single case research (SCR) data is necessary. Parker and Vannest (2009) emphasizes how the quantification of data is most needed when statements about the amount of improvement is necessary, and highlights documenting evidence of clinical effectiveness for governments or insurance purposes, and comparing the relative effectiveness of two or more interventions as examples. The necessity of quantifying data from SCR has recently prompted the development of several methods for calculating SCR effect sizes, and sparked a debate of which are most appropriate. The debate is usually been divided between those that believe regression-based approaches are most efficient and those that believe that non-regression-based approaches best quantify the effects of SCR (Parker et al., 2007). The regression-based approach recommends utilizing R^2, which is the most used regression-based effect size in all social science research, thereby giving it the advantage of being a well-known entity across several fields (Parker et al., 2007). It also has the advantage of easy conversion to Cohen's d, another popular effect size, as well as the capability to calculate confidence intervals and the utilization of all the data in all the phases. However there are some limitations connected to the use of regression-based analysis with SCR data. Parker and Vannest (2009) highlights how SCR data often fail to meet the parametric assumptions of serial independence, normality, and constant variance of residual scores. They also emphasize how regression-based effect sizes are disproportionately influenced by extreme outliers, which are common in SCR, and how interpretation of the common regression-based effect sizes usually are alien to visual analysis procedures (Parker & Vannest, 2009).

The non-regression-based approaches provide alternative methods to calculating the index of effect for SCR. Scruggs, Mastropieri, and Casto (1987) introduced percentage of non-overlapping data (PND) as one of the first non-regression-based method for calculating the effect size of SCR. PND is the percentage of Phase B data that are more extreme in the improved direction than the single most extreme Phase A data point. PND can be hand-calculated from a printed graph simply by drawing a line from the most extreme Phase A data point through the following treatment phase, before calculating the amount of Phase B data points that exceeds the line (Scruggs & Mastropieri, 1994). Scruggs and Mastropieri (1994) also suggested general interpretation guidelines for PND scores, where PND scores larger than 70% should be regarded as effective, PND scores between 50 and 70% as questionable, and studies with PND scores less than 50% should be regarded as not having an observed effect (Scruggs & Mastropieri, 1994). The PND method's strengths, as highlighted by Parker et al. (2007), can be viewed as its ease of calculation and applicability to any single subject design as well as its similarity to visual analysis. However Parker et al. (2007) also pointed out several weaknesses with the PND method. They highlighted how PND is not related to an established effect size, making individual interpretation guidelines necessary. They also specified that PND lack a known sampling distribution, thus making it impossible to calculate p values and confidence intervals. They point out how PND only include the most extreme data point from the A phase which, by definition, is the less reliable data point. They also highlight how PND may lack the ability to discriminate very successful interventions, and how it does not control for baseline trends (Parker et al., 2007).

In an effort to introduce an improved non-regression based approach to quantify SCR Parker et al. (2007) suggested percentage of all non-overlapping data (PAND) as an alternative to PND. Similarly to PND, PAND reflects non-overlap between phases, but differ in some important aspects (Parker et al., 2007). PAND incorporates all the overlapping data points from both phases, thereby avoiding a limitation of PND as well as allowing easy conversation to the already well established effect sizes phi and phi^2. As a result of the fact that phi and phi^2 have known samplings distributions, p values are accordingly attainable, statistical power can be calculated, and confidence intervals can be included to indicate the reliability of the effect size, thereby sidestepping another crucial limitation of the PND method. However the PAND method is still vulnerable to the limitation of insensitivity to the differentiation between the most successful interventions, as both PND and PAND result in a perfect 100% score if there is no data-overlap, regardless of the distance between the data points. PAND also share the limitation of not controlling for baseline trends with PND, thereby making it insensitive to the detection of potential changes during baseline as long as there is no overlap with the following intervention phase.

Parker and Vannest (2009) then introduced nonoverlap of all pairs (NAP) as a new and improved overlap-based effect size, with essential advantages over the PND and PAND. NAP is an index of overlap between phases in SCR. However it is distinguished by the fact that it compares all the data points in each phase with each other. NAP was introduced as a novel application of the effect size known mainly from medical diagnostic studies as Area Under the Curve (AUC) (Parker & Vannest, 2009). It can be calculated by hand by dividing the number of overlapping pairs with the number of total possibility of pairs. This can be accomplished by either simply counting all non-overlapping pairs or by counting all overlapping pairs and subtract from the total possibility pairs to obtain the non-overlap count, which is usually faster. NAP is therefor often calculated by giving each overlapping data point the score of 1, and each tie the score of .5. The overlap score is then added together before it is subtracted from the total number of pairs, which is calculated by multiplying the number of phase A data points with the number of Phase B data points. NAP can then finally be located by dividing the overlap count with the total possible pairs (Parker & Vannest, 2009). To illustrate this with an example: Imagine a simple AB graph with six Phase A data points and eight Phase B data points. Two of the six Phase A data points overlap with data points from the other phase. The first is a complete overlap with two of the Phase B data points, and a tie with one, giving it an overlap count of 2,5. The second overlapping Phase A data point is a complete overlap with one Phase B data point and a tie with another, thereby giving it the overlap score of 1,5. The number of overlapping pairs are then added together (2,5+1,5 = 4) and subtracted from the total possible pairs (48-4 =44), resulting in the sum of non-overlapping pairs (44). Finally the non-overlapping pairs are divided by the total possible pairs (44/48 = 0.9167), resulting in a NAP of .9167. Thus NAP is, in contrast to the other non-regression based methods, a result of the comparison of every single Phase A data point with each data point from the other phase.

Parker and Vannest (2009) displayed how NAP discriminated better among results from a large group of published studies, compared to the other non-overlap indices. They also highlighted that NAP should be less vulnerable to human error in calculation, due to the fact that NAP can be calculated by computerized methods**,** as AUC from a receiver operator characteristics curve (ROC) analysis. Another NAP advantage is that it, due to its large number of data comparisons, is the overlap-based effect size that has the strongest correlation with the widely used R^2 (Parker & Vannest, 2009). Parker and Vannest (2009) also displayed a strong correlation between NAP and visual judgments (.84) that distinguished it from other non-regressions based effect sizes.

The purpose of the present study was to review the evidence for tablet-based technology in the treatment of individuals with disabilities. In an effort to avoid the limitations connected to the use of regressions-based techniques with SCR, a non-regression based method was used. The NAP method was utilized, as it seems to have some notable advantages over the previous non-regressions based methods and is the overlap-based effect size with the strongest correlation to other established effect sizes.

Method

Study identification procedure

The studies included in this study were located through a search of the ProQuest and ERIC databases using combinations of the keywords: Autism, Tablet, and Ipad. In the search of the ProQuest database, it was also required that the keyword “tablet” was found in the abstract. We also included articles that were found through browsing the references section of selected studies. This search resulted in a total of 95 articles that were then evaluated for eligibility in the review.

Inclusion and exclusion criteria

The following six inclusion criteria were used to determine if studies were eligible to be included in the analysis. First, the study had to be published after 2010. Second, the study needed to employ a single-subject design. Third, the study had to provide empirical data. Fourth, the study was required to present the data graphically. Fifth, the study had to involve participants with ASD and/or intellectual disabilities (ID). Sixth, the study needed to use a tablet-based intervention. We used the following exclusion criteria. First, the tablet was used as a speak generating device (SGD), or second, the tablet was used for video modeling. Based on these criteria 15 of the articles were included in the analysis. The flowchart of the selection is displayed in Table 1. The studies evaluating the tablet used as a SGD were not deemed relevant for this study, as their effectiveness already have been thoroughly evaluated elsewhere (Alzrayer, Banda, & Koul, 2014; Kagohara et al., 2010; Lorah, Parnell, Whitby, & Hantula, 2015; Lorah et al., 2013; Roche et al., 2014; Stuart, 2012). Furthermore, in these studies the tablet is used for communication and not as an integrated component of the intervention. Studies using the tablet technology in video modeling procedures were also not considered applicable in this context, as video modeling already has been a subject for a large number of studies (Burke, Andersen, Bowen, Howard, & Allen, 2010; Cruz-Torres, 2015; Hughes & Yakubova, 2016; Plavnick, 2012; Weng, Savage, & Bouck, 2014) and the recently novel inclusion of tablets as the mediator was not regarded as sufficiently significant to justify an new analysis. Other studies were not included on a basis of being review articles with a lack of graphical representation of data (12), participants without developmental disorders (3), or the lack of tablet-based interventions (5).

The agreement in study inclusion was assessed. This was calculated by having two independent assessors rate if the 15 relevant studies would pass the inclusion criteria or not, and then calculated by dividing the number of agreements with the total number of agreements plus disagreements, multiplied with 100. The inter observer agreement (IOA) on the inclusion of the 15 studies were calculated to 100%, as both assessors agreed on all of the studies.

Study Coding Procedures

The 15 included studies were reviewed to assess the methodological rigor and evidential strength. The studies included in the review involved a total of 53 unique participants. The relevant studies were coded for the participant's age and diagnosis, the experimental design, design standards, the number off participants, the relevant skill category, setting, applied platform, and application name.

To evaluate the studies design standards the guidelines from Kratochwill et al. (2013) was adopted. Kratochwill et al. (2013) provided an overview of the What Works

Clearinghouse (WWC) standards for evaluating singe-case intervention research. They presented four criteria used to assess whether study's designs a) *meets standards*, b) *meets standards with reservation*, or c) *does not meet design standards*. Their first requirement is that an independent variable must be systematically manipulated. If this standard is not met, the study *does not meet design standards*. Secondly, each outcome variable must be measured systematically over time by more than one assessor. Minimum accepted values of interobserver agreement range from .80 to .90, if measured by percentage agreement and at least .60 if measured by Cohen's kappa. A summary of interobserver agreement for a variable must be based on at least 20% of the data points within each condition. If this standard is not met, the study *does not meet design standards*. Third, the study must include at least three attempts to demonstrate an intervention effect each at a different point in time. If this procedure is not followed, the study *does not meet design standards*. Their fourth criteria was that for a phase to qualify as an attempt to demonstrate an effect, the phase must include a minimum of three data points, with a preference for at least five (Kratochwill et al., 2013). Kratochwill et al. (2013) then provided specified requirements for the most common SCR designs. See Table 2 for the specific requirements used to evaluate the studies design standards. Specific requirements for multiple probes designs have not yet been developed, thus these studies were evaluated by the standards developed in relation to the multiple baseline design.

Analysis

Each study was examined for treatment efficacy through the calculation of NAP scores. The NAP scores of the relevant studies were obtained as direct output from raw scores as area under the curve from a receiver operator characteristic (ROC) diagnostic test module. The raw data scores were obtained from the respective studies through PlotDigitizer (Huwaldt & Steinhorst, 2013), a computer program that translate screenshots of graphs into raw data.

The raw data can then be used for an ROC analysis. Maintenance and generalization data was not included in the ROC analysis when available, as they were not considered directly relevant to the present study. Studies that employed reversal designs were calculated by comparing each treatment phase with the preceding baseline phase. Studies that used variations of alternative treatment designs were calculated by comparing the tablet-related treatment with the alternative treatments, usually consisting of traditional teacher-instructed interventions. Studies with multiple graphs were evaluated by analyzing the individual graphs independently, before they were added together through the ROC analysis, and NAP scores were then calculated representing the entire study. Studies that deployed different tablet applications were evaluated as if each application was done as a separate study, thereby resulting in individual NAP scores for each application.

Characteristics of included studies

Table 3 displays basic information for each application included in the analysis. This includes information regarding study design, design standards, participants (number, age and diagnosis), settings, skill domain, platform, as well as NAP score and P values. 11 of the studies used Ipad as the platform (73,33%), two studies used the Samsung galaxy note (13,33%), one study deployed HTML5 (6,67%), and one study used Smart table (6,67%). Five of the studies (33,3%) employed multiple probe designs, four studies (26,6%) used multiple baseline designs, three studies (20%) used variations of reversal designs, while two (13,3%) chose alternating treatment designs, and one (6,6%) used a comparison design. Most of the participants included in the relevant studies were at school-age level. The average age of the participant's was 9,5 with a range in age from 3 to 32 and a median of 11. Most of the studies were conducted in a school setting (66,67%), but the tablet-based applications were also implemented in clinical settings (13,33%), at home (6,67%), both at home and in the clinic (6,67%), as well as in a workplace (6,67%).

There was a diverse range of disabilities included in the studies: Autism spectrum disorder (77,4%); Intellectual disabilities (11,3%); Pervasive development disorder-not otherwise specified (9,4%); and Attention deficit hyperactive disorder (1,9%). The studies included in the analysis were categorized into skill domains on the basis of target behavior. Eight studies (53,3%) targeted academic skills, while five (33,3%) focused on social skills and two (13,3%) prioritized day-to-day skills.

Reliability

Reliability for NAP scores were assessed by having two independent researchers run the ROC analysis, before calculating NAP scores and comparing the results. Agreement was recorded when both raters reached the same NAP score from each study. The reliability score was then calculated by dividing the number of agreements with the total number of agreements plus disagreements, multiplied with 100. This calculation was completed for all of the included studies**.** Interobserver agreement (IOA) was also calculated for the evaluation of design standards. Two assessors independently evaluated the relevant studies and rated them with either *meets standards*, b*) meets standards with reservation,* or c) *does not meet design standards*, based on the criterions from Kratochwill et al. (2013). IOA was then calculated by dividing the number of agreements with the total number of agreements plus disagreements, multiplied with 100. The two independent assessors agreed on 19 of the NAP scores, and disagreed on two. Hence, the interobserver agreement of the study's NAP calculation was determined to be 90,48%. The two independent assessors agreed on all of the studies design standards, so the interobserver agreement for the designs standards were therefor calculated to 100%.

Results

All of the 15 studies focused on establishing or improving appropriate behavior through tablet-based interventions. Five studies, with a total of 12 participants, focused on improving social skills: eye contact (Jeffries, Crosland, & Miltenberger, 2016); turn-taking behaviors (Kim & Clarke, 2015); conversation skills (Sng, Carter, & Stephenson, 2017; Travers & Fefer, 2017); and play-behavior (Murdock, Ganz, & Crittendon, 2013). The average NAP score for these studies was calculated to .714 (SD = .2554). The NAP scores for the individual studies ranged from .4133 to .9897.

Eight studies targeted academic skills: Task completion (O'malley, Lewis, & Donehower, 2013); shared stories (Spooner, Ahlgrim-Delzell, Kemp-Inman, & Wood, 2014); science (Smith, Spooner, & Wood, 2013); on-task behavior (Arthanat, Curtin, & Knotak, 2013; Neely, Rispoli, Camargo, Davis, & Boles, 2013; Vandermeer, Beamish, Milford, & Lang, 2015); spelling (Seok, DaCosta, & Yu, 2015); and receptive identification (Chebli, Lanovaz, & Dufour, 2017). The average NAP score for all the studies aimed at improving academic skills was calculated to .7414 (SD = .2735), with a range within individual studies from .2500 to 1.

Two studies focused on improving day-to-day skills: sequencing skills (Doenyas, Şimdi, Özcan, Çataltepe, & Birkan, 2014); and café waitering (Cavkaytar, acungil, & Tomris, 2017). The average NAP for these studies were determined to .7955 (SD =2893), with individual studies ranging between 0.5909 and 1. The NAP score for all the 15 included studies was calculated to .7401 (SD = .2576), with a range of the individual studies from .2500 to 1. Table 4 exhibits the average NAP scores for each separate skill domain.

Discussion

The purpose of this study was to evaluate the effectiveness of interventions employing tablet-based technology in the treatment of children with autism and other disabilities. This was achieved by calculating the effect sizes of these studies through the NAP method. Several findings are evident from this review. One of the most important findings from this review is that the average NAP of .7401 indicates that the effect size of tablet-based interventions for individuals with developmental disabilities can be regarded as medium. However the range of individual NAP scores from .2500 to 1 indicate that not all tablet-based interventions are effective. Nine of the applications included in the analysis received NAP scores that indicate weak effects according to the guidelines from Parker and Vannest (2009). Parker and Vannest (2009) pointed out how NAP and the other non-overlap indices need their own interpretation guidelines and suggested that NAP scores of 0-.65 should be considered as weak effects, .66-.92 as medium effects, and .92-1,0 as large or strong effects. The relative high score necessary to attain the strong effects was intended as a counter measure to the limitation of non-overlap effect sizes in regards to discriminating between the most effective interventions (Parker & Vannest, 2009).

The large range of NAP scores illustrates that not all tablet-based interventions are effective. Even though tablets-based interventions often have potential advantages over other traditional formats, it does not automatically lead to effective interventions. Two of the studies included in this analysis provide good examples of this. Jeffries et al. (2016) attempted to establish eye contact with participants diagnosed with autism through a tablet-based interactive game. Their results showed that although all of the participants used the application successfully and requested to continue to use the tablet after completion, none of the participants displayed improved eye contact. However, a differential reinforcement procedure improved eye contact substantially with all of the participants. This highlights the importance of evaluating the effectiveness of tablet applications aimed at improving certain skills, and not only focusing on the general advantages of introducing tablets in interventions. The results from Chebli et al. (2017) contributes to these findings. They used a tablet-based intervention to teach one-word receptive identification to five children with ASD and test for generalization of the skills. Their results show that only three of the five participants displayed generalization, thereby underlining that tablet-based interventions might not be a one-size-fits all solution. There are several possible explanations to why tablet based interventions might be effective with some participants and not with others. One possible factor can be the respective participants prior experience with tablets. Some participant could have an extensive experience with tablets, which can influence the results in different directions. Prior experience with tablets in the context of leisure might result in the occurrence of problem behavior when there are demands and restrictions presented with the tablet. Prior experience with tablets in academic settings could however influence the speed of learning and reduce the tablet acclimation period. Future research should therefor consider further investigation of factors that influence the effectiveness and the degree of generalization from tablet-based interventions. The results from this analysis also indicate that the level of functioning of the participants might affect the effectivness and the degree of generalization obtained from tablet-based interventions. This highlights how future research should not only continue to investigate the effectiveness of tablets, but that it is also a necessity to further investigate the generalization value of tablet-based interventions.

The results from this analysis also indicate that tablets can be used with a variety of skills. The skills targeted in this analysis were divided in to three categories; academic, social, and day-to-day skills. All of the categories received a similar NAP score, indicating a medium effect. The fact that there was no significant discrepancy in the efficacy of these different skill categories demonstrates the tablets multipurpose ability. These results are similar to the results of Allen et al. (2016) and can be considered an important advantage for tablet-based interventions, as previous teaching equipment often have been single purpose and expensive.

Another significant finding in this study is that there was no overlap of applications employed. This can be a result of the sheer amount of applications available, which means that researchers are less likely to choose the same applications even though they focus on similar skills. Another possible factor is the fact that many researchers created their own applications and tailored them specifically to the participants and target behaviors in hand. This large variation of applications employed by researchers results in little cumulative empirical evidence for each application, which in turn means that caregivers and professionals are left with little guidance when it comes to choosing applications to implement in their training regimes. Another possible limitation that results from the manufacturing of applications especially for the purpose of conducting a study is that it might be exclusively customized to the particular participants in the study, and therefore lead to misleadingly positive results. This can result in studies that provide findings that are not generalizable to a larger sample of participants. Future research should target the evaluation of existing applications, in an effort to produce data that can be more informative for caregivers and professionals.

The results from this study indicate thus that tablet-based interventions are applicable to a large spectrum of skills, with similar effectiveness. This is especially notable in the interest of cost effectiveness. While a tablet can be considered costly to acquire, it can be cost effective in the long run, as it can be used for a multitude of different skills and situations, avoiding the need of different appliances for every skill, a weakness often seen prior in apparatuses. Tablets are thus a feasible and compelling option not only for caregivers, but also for schools or clinics that work with multiple children. The fact that there was considerable discrepancy in the skills being thought, as well in the techniques used in this study indicate that tablet applications often are highly customizable. Most of the applications included in the analysis were used with multiple participants, some with large varieties in age and diagnosis, which indicate that these applications were customizable and able to provide god results with variations of participants. Due to the fact that teaching based in applied behavior analysis requires highly individualized training procedures, it is most likely a necessity that applications are customizable. Future research should consider evaluating the degree of customizability of applications and compare that to the effectiveness of the application.

Even though most of the studies included in this analysis were conducted in a school setting, training was also done I a number of other settings, suggesting that tablet-based training can be done in multiple locations. This can be considered a notable results as it indicates that training can occur outside of the clinic and schools, and thereby allow for more training in the students home and other settings. This could have significant impact in EIBI programs, which require many hours of weekly training. Future research should therefore consider conducting more tablet-based interventions outside schools and clinics, to better understand the possibility of tablet-based interventions.

The results from this analysis also indicate that future research should further investigate and compare the difference in effectiveness between traditional teacher-instructed training and tablet-based training. Jeffries et al. (2016) provided some information in this regard, by demonstrating that teacher instructed differential reinforcement procedure were better at establishing eye contact with children with ASD than a tablet-based application. Even though it, in some instances, might be caused by poor applications, the difference in the rate of acquisition has yet to be thoroughly investigated. This could lead to a better understanding of the potential benefits and limitations with the separate training methods. It may, for instance, be the case that tablet-based instruction can require more trials to reach mastery, as it often lacks the customizability and sensitivity of teacher instructed training.

Some might argue that a limitation with this study is that it gives an understated impression of the effectiveness of the use tablets with children with developmental disabilities, due to the fact that studies employing tablets as speak generating devices or as facilitators for video modeling procedures have been excluded. Although this argument could be valid, indeed the average NAP scores would likely be higher if those studies were included, it can be argued that tablets as SGD or as facilitators of video models does not significantly contributed to those procedures, and would therefore lead to a falsely high NAP score when it comes to the evaluation of tablet-based interventions. When it comes to tablets as speak generating devices, versatile tablets, like the Ipad and the Galaxy Note, are mostly improved compared to prior SGD's in the regard that they are multipurpose, and therefore usable with different skills and tasks in addition to the SGD. This was however not regarded as sufficiently consequential to validate the inclusion in this analysis, as the tablet component was not deemed a significant factor in the interventions. Studies that utilized tablets as facilitators for video modeling were also excluded from the analysis due to the fact that they were not considered as employing the available technology sufficiently. The tablets were assessed as only contributing with the convenience of being portable compared to traditional video modeling procedures.

There are some general limitations with regard to the use of non-overlap effect sizes. Even though NAP displays a well-shaped uniform frequency it still exhibits a noticeable ceiling effect at the 80^{th} percentile. This is, however, a limitation with all non-overlap based indices and is only avoided by the regression-based methods that are able to measure degree of score separation beyond the complete overlap. Nonetheless the distribution of NAP has, together with PAND, displayed far less deficiencies compared to the other non-overlap indices, which has showed an incapability to discriminate among nearly half of datasets (Parker & Vannest, 2009). Another possible limitation with the NAP method is the concern with a lack of independence within time series data. The serial dependence in SCR is acknowledged to exist, but is mostly regarded to have little impact on effect sizes. In addition the possible drawback is best contained by the utilization of non-regression based techniques, which is less affected by the serial dependence and the benefits of applying statistical analysis to SCR are commonly viewed to outweigh the concerns (Parker & Vannest, 2009).

Another possible limitation of this study is the fact that studies employing a multiple probe design were evaluated by the criteria developed for multiple baseline design in regards to design standards. This might have resulted in these studies receiving a worse design standard grade then they were entitled to, due to not meeting the required amount of data points within each phase. It was, however, deemed the best options available, as there has not yet been presented any specific requirements to evaluate the design standards of multiple probe designs. Developing design standards for multiple probe designs should however be a target for future research, as it is a commonly used design in single case research.

The results from this analysis indicate that tablet-based interventions, on average, have a medium effect size. Regarding the relative novelty of tablet-based interventions, moderate effects are promising and indicate that its use should be further explored in both research and applied settings. However, deeper analyses uncover a large variety in effectiveness across the studies. These results are consistent with previous research and indicate that while tablet-based interventions often are effective, it is not a one-size-fits-all solution. Future research should therefor further investigate which factors that correlate with ineffective tablet-based interventions, in an effort to understand why it does not produce the same results with all participants. Even though the results from this study show that there was little difference in the effectiveness of the applications based on skill domain, future research should also use already available applications in an effort to provide more evidence supporting the effectiveness of specific applications as well as further investigate the generalization value of tablet-based interventions.

References

Allen, M. L., Hartley, C., & Cain, K. (2016). iPads and the use of "Apps" by children with Autism Spectrum Disorder: do they promote learning? *Frontiers in psychology, 7*.

Alzrayer, N., Banda, D., & Koul, R. (2014). Use of Ipad/Ipods with Individuals with Autism and other Developmental Disabilities: A Meta-analysis of Communication Interventions. *Review Journal of Autism Development Disorder*(1), 179-1991. doi:DOI 10.1007/s40489-014-0018-5

Arthanat, S., Curtin, C., & Knotak, D. (2013). Comparative observations of learning engagement by students with developmental disabilities using an iPad and computer: A pilot study. *Assistive Technology, 25*(4), 204-213.

Artoni, S., Buzzo, M. C., Buzzi, M., Fenili, C., Leporini, B., Mencarini, S., & Senette, C. (2013). A Portable Application for Supporting ABA Intervention. *Journal of assistive technologies, 7*(2), 78-92.

Ayres, K. M., & Langone, J. (2002). Acquisition and Generalization of Purchasing Skills Using a Video Enhanced Computer-based Intructional Program *Journal of Special Education Technology, 17*(4), 15-28.

Baer, D. M. (1977). "Perhaps it would be better not to know everything". *Journal of applied behavior analysis, 10*(1), 167-172.

Basil, C., & Reyes, S. (2003). Acquisition of literacy skills by children with severe disability. *Child Language Teaching and Therapy, 19*(1), 27-48.

Burke, R. V., Andersen, M. N., Bowen, S. L., Howard, M. R., & Allen, K. D. (2010). Evaluation of two instruction methods to increase employment options for young adults with autism spectrum disorders. *Research in developmental disabilities, 31*(6), 1223-1233.

Cavkaytar, A., acungil, A. T., & Tomris, G. (2017). Effectiveness of Teaching Café Waitering to Adults With Intellectual Disability through Audio-Visual Technologies. *Education and Training in Autism and Developmental Disabilities, 52*(1), 77-90.

Chebli, S. S., Lanovaz, M. J., & Dufour, M.-M. (2017). Generalization Following Tablet-Based Instruction in Children With Autism Spectrum Disorders. *Journal of Special Education Technology, 32*(2), 70-79.

Cruz-Torres, E. M. (2015). *Promoting daily living skills for adolescents with autism spectrum disorders via parent delivery of video prompting on the iPad*: Florida Atlantic University.

Doenyas, C., Şimdi, E., Özcan, E. Ç., Çataltepe, Z., & Birkan, B. (2014). Autism and tablet computers in Turkey: Teaching picture sequencing skills via a web-based iPad application. *International Journal of Child-Computer Interaction, 2*(1), 60-71.

Egeberg, G., Hultin, H., & Berge, O. (2017). *Monitor skole 2016: skolens digitale tilstand* (2 ed.): Senter for IKT i utdanningen.

El Zein, F., Gevarter, C., Bryant, B., Son, S.-H., Bryant, D., Kim, M., & Solis, M. (2016). A comparison between iPad-Assisted and teacher-directed reading instruction for students with Autism Spectrum Disorder (ASD). *Journal of Developmental and Physical Disabilities, 28*(2), 195-215.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2010). Using participant data to extend the evidence base for intensive behavioral intervention for children with autism. *American Journal on Intellectual and Developmental Disabilities, 115*(5), 381-405.

Green, G. (2011). Early Intensive Behavior Analytic Intervention for Autism Spectrum Disorders. In E. A. Mayville & J. A. Mulick (Eds.), *Behavioral Foundations of*

Effective Autism Treatment (pp. 183-199). Cornwaal-on-Hudson, NY: Sloan Publishing.

Hetzroni, O., Rubin, C., & Konkol, O. (2002). The use of assistive technology for symbol identification by children with Rett syndrome. *Journal of Intellectual and Developmental Disability, 27*(1), 57-71.

Hughes, E. M., & Yakubova, G. (2016). Developing Handheld Video Intervention for Students With Autism Spectrum Disorder. *Intervention in School and Clinic, 52*(2), 115-121.

Huwaldt, J. A., & Steinhorst, S. (2013). Plot digitizer.

Jeffries, T., Crosland, K., & Miltenberger, R. (2016). Evaluating a tablet application and differential reinforcement to increase eye contact in children with autism. *Journal of applied behavior analysis, 49*(1), 182-187.

Kagohara, D. M., van der Meer, L., Achmadi, D., Green, V. A., O'Reilly, M. F., Mulloy, A., . . . Sigafoos, J. (2010). Behavioral intervention promotes successful use of an iPod-based communication device by an adolescent with autism. *Clinical Case Studies, 9*(5), 328-338.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*: Oxford University Press.

Kim, S., & Clarke, E. (2015). Case study: An iPad-based intervention on turn-taking behaviors in preschoolers with autism. *Behavioral Development Bulletin, 20*(2), 253.

Knight, V., McKissick, B., R. , & Saunders, A. (2013). A Review of Technology-Based Interventions to Teach Academic Skills to Students with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 43*, 2628-2648. doi:10.1007/s10803-013-1814-y

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education, 34*(1), 26-38.

Lorah, E. R., Parnell, A., Whitby, P. S., & Hantula, D. (2015). A systematic review of tablet computers and portable media players as speech generating devices for individuals with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(12), 3792-3804.

Lorah, E. R., Tincani, M., Dodge, J., Gilroy, S., Hickey, A., & Hantula, D. (2013). Evaluating picture exchange and the iPad™ as a speech generating device to teach communication to young children with autism. *Journal of Developmental and Physical Disabilities, 25*(6), 637-649.

Murdock, L. C., Ganz, J., & Crittendon, J. (2013). Use of an iPad play story to increase play dialogue of preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*(9), 2174-2189.

Neely, L., Rispoli, M., Camargo, S., Davis, H., & Boles, M. (2013). The effect of instructional use of an iPad® on challenging behavior and academic engagement for two students with autism. *Research in Autism Spectrum Disorders, 7*(4), 509-516.

O'malley, P., Lewis, M., & Donehower, C. (2013). Using tablet computers as instructional tools to increase task completion by students with autism. *Online submission*.

Parker, R. I., Cryer, J., & Byrns, G. (2006). Controlling baseline trend in single-case research. *School Psychology Quarterly, 21*(4), 418.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND) an alternative to PND. *The Journal of Special Education, 40*(4), 194-204.

Parker, R. I., & Vannest, K. (2009). An Improved Effect Size for Single-Case Research: Nonoverlap of All Pairs. *Behavior Therapy, 40*, 357-367.

Plavnick, J. B. (2012). A practical strategy for teaching a child with autism to attend to and imitate a portable video model. *Research and Practice for Persons with Severe Disabilities, 37*(4), 263-270.

Roche, L., Sigafoos, J., Lancioni, G. E., O'Reilly, M. F., Schlosser, R. W., Stevens, M., . . . James, R. (2014). An evaluation of speech production in two boys with neurodevelopmental disorders who received communication intervention with a speech-generating device. *International Journal of Developmental Neuroscience, 38*, 10-16.

Scruggs, T. E., & Mastropieri, M. A. (1994). The utility of the PND statistic: A reply to Allison and Gorman. *Behaviour Research and Therapy, 32*(8), 879-883.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24-33.

Seok, S., DaCosta, B., & Yu, B. M. (2015). Spelling practice intervention: A comparison of tablet PC and picture cards as spelling practice methods for students with developmental disabilities. *Education and Training in Autism and Developmental Disabilities, 50*(1), 84.

Shimizu, H., Twyman, J. S., & Yamamoto, J. (2003). Computer-based sorting-to-matching in identity matching for young children with developmental disabilities. *Research in developmental disabilities, 24*(1), 183-194.

Simpson, A., Langone, J., & Ayres, K. M. (2004). Embedded video and computer based instruction to improve social skills for students with autism. *Education and Training in Developmental Disabilities*, 240-252.

Smith, B. R., Spooner, F., & Wood, C. L. (2013). Using embedded computer-assisted explicit instruction to teach science to students with autism spectrum disorder. *Research in Autism Spectrum Disorders, 7*(3), 433-443.

Sng, C. Y., Carter, M., & Stephenson, J. (2017). Teaching a Student With Autism Spectrum Disorder On-Topic Conversational Responses With an iPad: A Pilot Study. *Australasian Journal of Special Education, 41*(1), 18-34.

Spooner, F., Ahlgrim-Delzell, L., Kemp-Inman, A., & Wood, L. A. (2014). Using an iPad2® with systematic instruction to teach shared stories for elementary-aged students with autism. *Research and Practice for Persons with Severe Disabilities, 39*(1), 30-46.

Stuart, S. (2012). Finding a Voice. *Principal, 91*(4), 32.

Tincani, M., & Boutot, E. (2005). Technology and autism: Current practices and future directions. *Handbook of special education technology research and practice. Whitefish Bay, WI: Knowledge by Design*, 413-421.

Travers, J. C., & Fefer, S. A. (2017). Effects of Shared Active Surface Technology on the Communication and Speech of Two Preschool Children With Disabilities. *Focus on Autism and Other Developmental Disabilities, 32*(1), 44-54.

Vandermeer, J., Beamish, W., Milford, T., & Lang, W. (2015). iPad-presented social stories for young children with autism. *Developmental neurorehabilitation, 18*(2), 75-81.

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical psychology review, 30*(4), 387-399.

Weng, P.-L., Savage, M. N., & Bouck, E. C. (2014). iDIY: Video-based instruction using iPads. *TEACHING Exceptional Children, 47*(1), 11-19.

Yamamoto, J.-i., & Miya, T. (1999). Acquisition and transfer of sentence construction in autistic students: Analysis by computer-based teaching. *Research in developmental disabilities, 20*(5), 355-377.

Tables

Table 1

The flowchart of the study's selection process

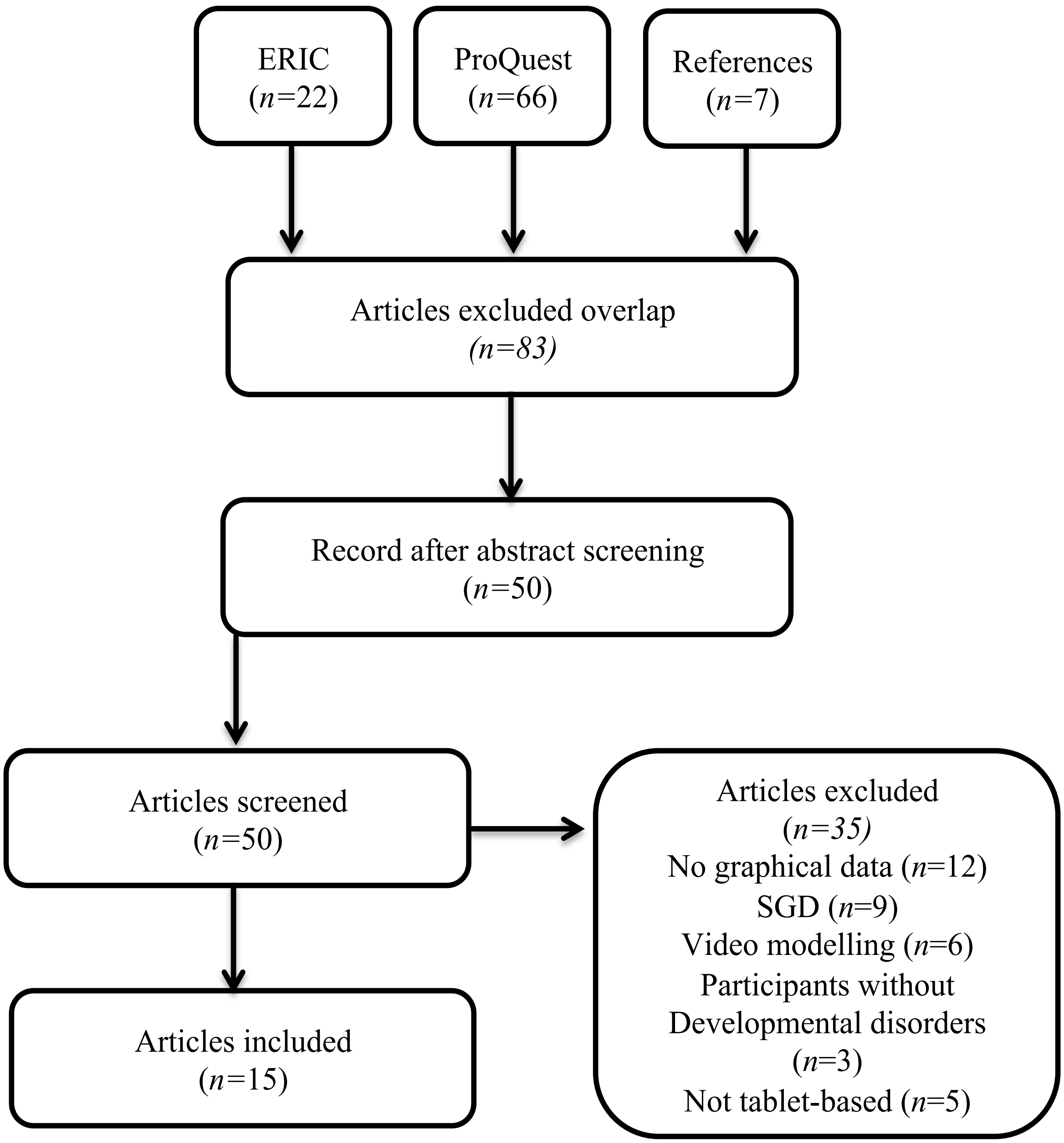

Table 2

The specific design standard requirements used to evaluate the studies design standards (Kratochwill et al., 2013).

Design	Meets standards	Meets standards with reservation	Does not meet design standards
Reversal	Min. four phases Min. five data points in each phase	Min. four phases Min. three data points in each phase	One or more phases Less than three data points
Multiple baseline	Min. six phases Min. five data points in each phase	Min. six phases Min. three or four data points in each phase	One or more phases Less than three data points
Alternating treatment	Min. five repetitions of the alternating sequence	Min. four repetitions of the alternating sequence	Less than four repetitions of the alternating sequence

Table 3

Summarization of important study features

Study	Design	Designstandard	*N*	Age (years)	Diagnosis	Skill	Setting	Platform	Application	NAP	P Value
Doenyas et al. (2014)	ABA	Meets standards	*3*	4, 11, 15	ASD	Day-to-day skills	School	HTML5	Custom*	.5909	>.001
Jefferies et al. (2016)	MB	Meets standards	*3*	3, 3, 5	ASD	Social	Clinic	Ipad	Look in My Eyes Steam Train	.4133	.0499
Kim and Clark (2015)	MB	Does not meet design standards	*2*	4, 4	ASD	Social	Home	Ipad	iTake turns & PowerPoint	.8896	>.001
Travers and Fefer (2017)	ATD	Meets standards	*2*	4, 5	ASD, ID	Social	School	Smart Table	Pre-installed drawing application	.4771	.0087
O'Malley et al. (2013)	ABAB	Meets standards with reservation	*7*	11 − 13	ASD	Academic	School	Ipad	Matching Game-My First Numbers app	.5476	.7639
Spooner et al. (2014)	MP	Meets standards	*4*	12, 8, 11, 8	ASD	Academic	School	Ipad2	GoTalk Now	.9828	>.001
Smith et al. (2012)	MP	Meets standards	*3*	12, 11, 12	ASD, ID, AD/HD	Academic	School	Ipad2	Keynote	1	>.001
Vandermeer et al. (2013)	MB	Meets standards with reservation	*3*	4, 4, 4	ASD	Academic	School	Ipad	Stories2Learn	.785	>.001
Murdock et al. (2013)	MB	Meets standards with reservation	*4*	4, 4, 4, 4	PDD-NOS, ASD	Social	Clinic	Ipad	Keynote	.8004	>.001
Neely et al. (2013)	ABAB	Meets standards	*1*	7	AS	Academic	Home	Ipad	WritePad	1	>.001
Neely et al. (2013)	ABAB	Meets standards	1	3	PDD-NOS	Academic	School	Ipad	Little Matchups	1	>.001
Seok et al. (2015)	ATD	Meets standards	3	8, 15, 16	ASD, TS, ID	Academic	School	Samsung Galaxy Note	Play with the Korean Language	.5439	>.001
Arthanat et al. (2013)	AB	Does not meet design standards	1	12	ASD	Academic	School	Ipad	ABC - Letters, Numbers, Shapes, and Colors	.2500	.5403
Arthanat et al. (2013)	AB	Does not meet design standards	1	12	ASD	Academic	School	Ipad	Abby Train Colors	.8438	.1124
Arthanat et al. (2013)	AB	Does not meet design standards	1	12	ASD	Academic	School	Ipad	Puzzle Math	1	.2207
Arthanat et al. (2013)	AB	Does not meet design standards	1	12	ASD	Academic	School	Ipad	Kids can Spell	1	.0209
Arthanat et al. (2013)	AB	Does not meet design standards	2	13, 11	ASD	Academic	School	Ipad	Jungle Coins	.3347	.2939
Arthanat et al. (2013)	AB	Does not meet design standards	2	13, 11	ASD	Academic	School	Ipad	Play words	.4831	.1298
Cavkaytar (2017)	MPD	Meets standards with reservation	3	32, 19, 26	ID	Day-to-day skills	Work	Ipad	Cafe Waiter Education Program	1	>.001
Sng et al. (2017)	MPD	Does not meet design standards	1	7	ASD	Social	School	Ipad	The Conversation Coach	.9897	>.001
Chebli (2017)	MPD	Does not meet design standards	5	10, 7, 4, 5, 11	ASD	Academic	School	Samsung Galaxy Note	OpenSource Discrete Trial Instructor	.6093	>.001

Note. ABAB = Reversal Design; MB = Multiple Baseline Design; ATD = Alternative Treatment Design; MP = Multiple Probe Design; AB = Comparison Design; ASD = Autism Spectrum Disorder; ID = Intellectual Disabilities; AD/HD = Attention Deficit Hyperactivity Disorder; PDD-NOS = Pervasive Developmental Disorder-Not Otherwise Specified; AS = Asperger Syndrome; TS = Turner Syndrome, NAP = Non Overlap of All Pairs.

Table 4

Average NAP scores for each skill domain

Skill	*n*	NAP	SD	Range
Social	*12*	.71	.26	.41-.99
Academic	*35*	.74	.27	.25-1
Day-to-day skills	*6*	.80	.29	.59-1
Total	53	.74	.26	.25-1

Establishment and Generalization of Receptive Labeling following Tablet-based Instruction in Children with Autism

Vetle Berge

Oslo Metropolitan University

Abstract

The use of tablet technology has become widespread in teaching based on applied behavior analysis. However, few studies have investigated the degree of establishment and generalization of skills following tablet-based interventions. Such an investigation is particularly urgent due to the significant difference between operating a tablet and interacting with others in a natural environment.. As a result of this, the current study aimed at teaching receptive language skills with five participants with ASD through the tablet application Superspeak and test for generalization to the natural environment. Only one of five participants learned receptive labels through the app. This participant generalized labeling to other exemplars of the same objects in the natural environment. These findings indicate that tablet-based interventions might work well for some, but not all children with autism. Possible reasons for this are discussed. Future research should further examine possible prerequisite skills that are needed to benefit from tablet-based teaching and how the applications are designed.

Keywords: Tablets, Technology, Generalization, Receptive language, Autism, Behavior Analysis

Establishment and Generalization of Receptive Labeling Following a Tablet-based Instruction in Children with Autism

Autism spectrum disorder (ASD) is now diagnosed in one out of every 68 children in the United States (Bhat, Acharya, Adeli, Bairy, & Adeli, 2014). In Norway the prevalence appears to be lower. A recent study found that ASD was diagnosed in one out of 189 children (Özerk, 2016) Özerk (2016) also highlighted how there has been a significant rise of prevalence rates of ASD over the last decades in Norway. This represents a significant challenge for health care and educational programs. ASD is a comprehensive diagnosis, and is recognized by severe impairment in social interactions and communication, as well as high rates of stereotypical behavior (American Psychiatric Association, 2013). Individuals diagnosed with ASD often require professional assistance throughout their lives, although early intensive behavioral intervention (EIBI) has been reported to significantly improve the rate of learning (Klintwall, 2015), as well as the intellectual and adaptive functioning (Eldevik et al., 2010). An important skill domain that often plays an imperative part in EIBI training is language. This is often a significant deficit seen with children with ASD, and is often one of the first problem areas recognized by caregivers (Lovaas, 2003). Language deficits are especially important as it ties into many other behavioral deficits and excesses. The occurrences of tantrums and self-injurious behaviors often serve as a form of as non-vocal communication based on an inadequate mastery of language, and improved language might contribute to reduce these kinds of behavioral excesses (Lovaas, 2003).

Language is often divided into receptive and expressive domains, generally initiated by the receptive component (Lovaas, 2003). Receptive language can be understood as language comprehension, while expressive language refers to speaking. Lovaas (2003) described the dichotomy between the two as how much individuals understand what is being said to them, referring to receptive, and how individuals express words, referring to expressive. Receptive language is often trained through the teacher giving the student a vocal instruction and the student responding by behaving accordingly. The stimulus is therefore vocal and the response is non-vocal (Lovaas, 2003). An example of a typical receptive language exercise may be that the teacher asks the student to touch a cup, and the student touches the cup. When the student's non-vocal response corresponds with the vocal instruction, one may infer that the student has at least acquired part of the meaning of the instruction (Lovaas, 2003). The simplicity of these procedures has led receptive language tasks to be especially suited to technology-based equipment.

The recent years advances in technology have led to an exceptional growth of the use of tablet applications in education. Clarke and Svanaes (2014) conducted a thorough literature review and concluded that the use of tablet technology in education is widespread. Handheld devices are not only frequently used at schools, but are also an important part of many students leisure time. A report conducted by the Norwegian Centre for Intelligence and Communication Technology in Education reported that 90,7% of Norwegian 7th graders have access to an handheld devices in their home, with 68,5% being in the possession of their own device (Egeberg, Hultin, & Berge, 2017). Recently applications used with these handheld devices have been designed for the treatment of individuals with developmental disabilities. The applications are often expensive and are regularly marketed towards parents and caregivers with the assumption that they are effective. An overview by Autism speaks, however, indicates that only a small number of these applications have had their efficacy scientifically tested. According to their summary only 1 out of 86 of the behavioral intervention applications have solid empirical evidence supporting their efficacy (Autism Speaks, 2017).

Some recent studies have however started to investigate the effectiveness of the larger term technology-based interventions, mainly including interventions based on computers.

Knight, McKissick, and Saunders (2013) conducted a comprehensive literature review to determine the efficacy of technology-based interventions to teach academic skills to individuals with ASD. They found that there were low levels of evidence for the use of technology-based interventions. They highlighted the need for further research and recommended that it should be used with caution. DiGennaro Reed, Hyman, and Hirst (2011) conducted a literature review on the implementation of technology to teach social skills to children with autism. They concluded that even though their results indicated that technology could be integrated into interventions targeting social skills deficits, there was little research investigating its efficacy.

However, there are some elements of the comprehensive treatment of young autistic children that has been proven efficient for the use of tablet technology. Tablet applications have been proven effective in augmentative and alternative communication training (AAC), where the tablet applications mainly function as a speech generating device (SGD) in requesting preferred stimuli, comparable to the function of pictures in the picture exchange communication system (PECS) (Achmadi et al., 2012; Flores et al., 2012; Kagohara et al., 2010; van der Meer et al., 2011). Other studies have proven that tablet applications can be effective as video modeling tools (Kagohara, Sigafoos, Achmadi, O'Reilly, & Lancioni, 2012) and as prompting tools for academic tasks for children with autism (Van Laarhoven, Johnson, Van Laarhoven-Myers, Grider, & Grider, 2009). These studies, as well as the large majority of published studies, focus on using tablets to mediate already established effective training procedures, or integrate a major human component. The integration of a major human component can often hinder the potential benefits from tablet-based teaching, as it would not be significantly different from traditional procedures. Bosseler and Massaro (2003) used for instance computer-based instruction (CBI) to teach vocabulary to children, but relied on edible reinforcers delivered by an instructor to maintain the target behavior. In a similar fashion Sansosti and Powell-Smith (2008) had to modify their procedure, with two out of three participants, to include teacher prompting and peer social reinforcement to increase social communication skills through computer-presented social stories. This human dependency appears to be common in CBI and tablet-based training, especially in regards to prompting and reinforcement (Burton, Anderson, Prater, & Dyches, 2013; Ganz, Boles, Goodwyn, & Flores, 2014; Jowett, Moore, & Anderson, 2012; Murdock, Ganz, & Crittendon, 2013; Schery & O'Connor, 1997; Still, May, Rehfeldt, Whelan, & Dymond, 2015).

However, few studies have investigated the effectiveness of CBI when the application itself does the teaching, thereby promoting autonomy and reducing teacher dependency. Applications performing the essential components of behavioral training could, if effective, have considerable impact, as they theoretically would allow training to occur in the absence of an instructor. Considering the high number of treatment hours recommended to ensure optimal outcome (Eldevik et al., 2010; Virués-Ortega, 2010) a reduction in the extent of time that a teacher is required would not only increase the autonomy of the student but also significantly reduce the cost of treatment (Shah, 2011) and make it accessible for a greater part of the population.

Another area lacking in the current research is the deliberate testing of generalization of the skills established with tablets. As elucidated by Stokes and Baer (1977) generalization cannot be considered a passive phenomenon that is expected to occur naturally. It is essential that generalization is considered an operant response that is programed and tested for (Stokes & Baer, 1977). Over a quarter-century after the classic article on generalization Osnes and Lieblein (2003) addressed the status of generalization-promotion in behavior analysis. They found that even though generalization requires much effort to be obtained, a considerable amount of research has been designed solely to demonstrate the functional relationship between training variables and generalization. Consequently, they concluded that the acknowledgment for the necessity of generalization promotion is a well-established part of behavior analysis that has resulted in a growing data base across diverse areas of the field (Osnes & Lieblein, 2003).

The novelty of tablet technology seems to have led the early research to concentrate on investigating its effectiveness, and not so much on evaluating the generalization effects of these interventions. This is a significant gap in the research considering that the nature of tablet-based teaching is very different from the natural environment. This is especially important considering the substantial amount of time and resources that are often devoted to these applications.

Jeffries, Crosland, and Miltenberger (2016) are among the few who have scientifically examined the grade of generality where the application was the major component in the training of a specific skill. They examined whether an application aimed at establishing eye contact, actually improved eye contact in a natural setting and compered the tablet training to a differential reinforcement procedure done by a teacher. Their results showed that even though all the participants successfully completed the training, and wanted to continue to use the application after completion, none of the children improved eye contact in the natural setting. Whereas, eye contact increased when a teacher used differential reinforcement procedures. This emphasizes the importance of assessing generalization effects of applications used in treatment of children with learning disabilities.

Another recent study by Chebli, Lanovaz, and Dufour (2017) targeted generalization following tablet based instruction with children with ASD. They employed a multiple probe design across concepts in an effort to evaluate the effectiveness of teaching one-word concepts, such as cow or dress, through the application OpenSource Discrete Trial Instructor. This application was developed by the authors, based on the principles of applied behavior analysis, and provided instruction in a discrete trial format. The application had a integrated prompt-procedure and video reinforcement. To test for generalization, the authors used both real-life objects and untaught images of the mastered objects. Their results showed that three of the five children showed generalization on at least two out of three concepts following tabled-based instruction. However, two of the five participants never displayed generalization despite completing more than 60 training sessions. These results underline again the necessity of assessing the generalization to real-life and indicate that tablet-based instruction may not be a one-size-fits all solution. This also shows the importance of identifying characteristics in the students that can predict the effectiveness of tablet-based interventions.

Together these studies provide an important first step towards providing a thorough scientific understanding of the generalization effects of tablet-based interventions. However, based on these mixed results, it is clear that this is an area in need of further investigation. The purpose of the present study was to assess whether receptive labeling could be established through the tablet application Superspeak and to assess generalization to other similar objects in natural settings.

Method

Participants

Five participants with developmental disabilities were recruited to the study. These participants were selected through cooperation with the Center of Early Intervention (STI) in Oslo. The participants were located in different schools and pre-schools around Oslo. To participate in the study the participants had to have a developmental disability diagnosis, a receptive vocabulary of between 25 and 100 words and prior experience with discrete trial training. These criteria were employed in an effort to locate participants who had a small, but existing, receptive vocabulary and had prior experience with the structural format of the training. The minimum receptive vocabulary requirement was set as it was deemed too time consuming to start receptive labeling training for children with no such skills. The maximum was employed to make sure that receptive labeling would be a socially significant skill for the participants and that the skill level would represent typical users of the application. The prior experience with discrete trial training was required to make sure the participants were familiar with the format of the training sessions. A total of nine participants were referred from STI, of which four were excluded because they did not meet the criteria for participation. The participants' teachers agreed to assist the study, and the parents of the participants signed a consent form. Pseudonyms have been used to protect the identity of the participants.

Ethan was 18-year-old boy with the diagnosis pervasive development disorder-not otherwise specified (PDD-NOS) and Down syndrome. He attended a school in Oslo for children with ASD. Ethan was reported to be visually competent and thrived with the use of pictures. He communicated through a combination of one-word sentences and PECS. He received one-on-one instruction from a teacher, and was regularly taught different skills in a discrete trial format.

Jonathan was a six-year-old boy diagnosed with ASD. He attended a regular preschool in Oslo and received individual instruction from a teacher in a separate room. Jonathan communicated through two-word-sentences and was reported to have a receptive vocabulary of approximately 100 words. He was regularly taught several skills in a discrete trail format.

Henrik was an eight-year-old boy diagnosed with ASD. He attended a primary school in Oslo. He communicated through a combination of one-word sentences and PECS. Henrik was reported to have a receptive vocabulary of between 50 and 100 words. He received one-on-one instruction from a teacher in a discrete trail format.

Billy was a twelve-year-old boy with a dual diagnosis of ASD and Down syndrome. He attended a primary school outside of Oslo and was placed in a department that specialized in the treatment of children with developmental disabilities. He was reported to have a receptive vocabulary of between 50 to 100 words. Billy communicated with a combination of one-word-sentences and PECS. He received one-on-one instruction from a teacher in a discrete trial format.

Simon was a 13-year-old boy with a dual diagnosis of ASD and Down syndrome. He attended the same school as Billy and was placed in the same department. Billy was reported to have a receptive vocabulary of approximately 50 words, and mainly communicated through sign language and PECS. He received one-on-one instruction in a discrete trial format.

Material and Setting

The training was implemented through the use of a tablet, more specifically an Ipad and the application Superspeak. This is an application developed specially for children with developmental disabilities and it is designed based on the principles from applied behavior analysis. The application provides instruction, reinforcement and error-correction procedures in an effort to keep teacher involvement to a minimum. The application includes a number of procedures to teach different skills, for instance matching and joint attention, but in this study it was only used to teach receptive labeling. The teaching was conducted in the participant's regular workspaces with a teacher present. The teacher involvement was kept to a minimum, and largely consisted of prompting the participants to attend to the tablet. The generalization trials were conducted with in-vivo objects of the pictures taught on the tablet. The testing was conducted with both the teacher and the first author present. The teacher conducted the generalization probes, with the first author recording the participants responding for the assessment of inter observer agreement (IOA). IOA was recorded for 100% of the generalization trials and was calculated using the following formula (agreements/agreements+disagreements)x100. IOA was calculated to 100%. The participants responding during tablet teaching was automatically recorded by the application and sent by e-mail to the first author.

Design

The intervention was implemented in a multiple probe design across different groups of stimuli. An assessment was completed prior to the start of the study in order to evaluate if the participants qualified for the study and to identify appropriate tasks for receptive labeling. The intervention was staggered across stimuli groups, and the procedure was replicated across participants in order to improve experimental control.

Procedure

Baseline probes was conducted to assess the performance level prior to the intervention in both the tablet condition and the tabletop condition with all stimuli groups. Baseline probes consisted of one programmed test session on the tablet. This session included nine trials were the participant did not receive any consequences contingent on performance. The participant received social reinforcement contingent on on-task behavior during the baseline probes. Generalization probes consisted of a similar session conducted on a tabletop setting with a teacher and real object exemplars of the pictures taught on the tablet. A teacher familiar with the participant conducted the generalization probes, while the first author observed and recorded the participants responding. The teachers were instructed to not practice any of the relevant stimuli groups outside of the project.

The intervention consisted of training receptive labels through the Ipad application Superspeak. Different stimuli groups were customized to each participant and programmed into separate games. Each game included one stimuli group with three different objects. One stimuli group was for instance animals, consisting of sheep, monkey and lion. The receptive labels of these objects where taught through establishing a relation between the name of the object and two different picture exemplars of each object. This was accomplished by separating the game into different steps. The first step included the relation between the vocal sound of the objects and the first set of object pictures. Once the participants had reached a mastery criterion of 90% correct responding within a block of nine trials, with three trials for each object, the game advanced to the next step. The second step of each game was a test including nine trials equal to the trials in step one, but without the delivery of feedback. Step three consisted of training the relation between the sounds of the object names to the second set of pictures. Once the mastery criterion was reached in this step, the game advanced to step four. This step was a test for the relation established in step three. Upon completion of this second test, step five was introduced. Step five implemented a mix of the two prior training steps, thereby mixing the two different pictures related to the same vocal instruction. Once the mastery criterion was reached within a block of nine trials the last step was introduced. This step was a final test, consisting of three trials on each of the relations trained in random order. Thereby making the final step amount to 18 trials. The tablet-based training procedure is displayed in table 1.

The participants trained with the application on two separate sessions a day, each lasting for approximately 15 minutes, five days a week until mastery. The training trials consisted of presentation of the vocal conditional stimulus "touch x" accompanied by a field of three pictures. The consequence for touching the corresponding picture was a "correct sound" and a green check mark appearing on the picture**.** The completion of one in-game block led to the removal of a puzzle piece blocking a pre-set picture. Upon finalization of the game all the pieces blocking the picture was removed thereby unveiling the complete picture. Touching the wrong picture led to an error-correction procedure consisting of four steps. The first step faded the color for the wrong options, brightened the color of the correct picture and presented a point model for choosing the right option. The second step also included color fading for the wrong options, but instead of a model, this step consisted of a shaking nudge of the correct picture. The third step in the error-correction procedure showed an enlargement of the correct option, and in the fourth and final step the stimuli was presented without any help.

If the participant responded incorrectly in the error-correction procedure, the preceding step in the procedure was re-introduced. The tablet-based error-correction procedure is displayed in table 2. If the participant did not respond within five seconds, a vocal prompting procedure was introduced. This procedure consisted of a trial with the instruction “touch the picture”. If the participant touched the picture the regular procedure continued. If the participant did not respond within five seconds the instruction was repeated. If the participant still did not respond, an incorrect response was registered and the error-correction procedure was introduced.

In accordance with the multiple probe design across stimuli groups, the intervention was introduced to the second group of stimuli when the first group was acquired, and to the third when the second was acquired. New baseline and generalization probes were conducted with each group prior to the introduction of the intervention. After the completion of a game the participants completed three post-training probes, equal to the baseline probes. The skills were then probed for generalization by completing three tests with the same procedure as the post-training test, but with a teacher in a tabletop setting with real objects. Participants were also tested for maintenance with the prior stimuli-sets after the completion of a new set.

Results

Ethan displayed low rates of correct responding during the pre-tests of the first stimuli group, with both the tablet (M = 22,2%) and the in-vivo objects (M = 33,3). In the post-test he demonstrated high rates of correct responding with the tablet (M =96.3) as well as the in-vivo objects (M =96.3). High rates of correct responding were also present during the maintenance condition with both the tablet (M = 94,4) and the in-vivo objects (M =88,8). Similar low levels of correct responding were demonstrated during the pre-tests of the second stimuli group, with the tablet (M =38,9) as well as with the in-vivo objects (M =44,4). The post-test results from the second stimuli group displayed similarly high levels with tablet (M =100%)

while the in-vivo objects showed slightly more variation then during the first stimuli group (*M* =81,5%). The maintenance condition demonstrated similar results as the preceding post-test, with high levels of correct responding with the tablet (*M* =100%) and slightly less with the in-vivo objects (*M* =83,3%). The third stimuli group displayed analogous levels of responding to the other stimuli groups during pre-tests, with the tablet (*M* =37,8%) as well as the in-vivo objects (*M* =44,4%). The post-test yielded high levels of correct responding with the tablet (*M* =100%), however the results from the in-vivo objects were lower then during any other condition (*M* =59,2%). The maintenance condition displayed lower levels of correct responding with tablet (77,2%) and higher levels with the in-vivo objects (77,2%) then in the post-test. Figure 1 displays the percentage of correct responding for all stimuli-groups during the pre-tests, post-tests, and maintenance conditions for Ethan

Henrik initially displayed moderate levels of correct responding before it gradually increased. He did not meet the mastery criteria in 500 trials during training. Figure 2 exhibit the percentage of correct responding and the number of trials during training of the first stimuli group with Henrik. Jonathan displayed moderate levels of responding throughout the training. Jonathan did not meet the mastery criteria in 228 trials under the course of the training. Figure 3 displays the percentage of correct responding and the number of trials during training of the first stimuli group with Jonathan. Billy demonstrated moderate levels of responding during the training. He did not reach the mastery criterion in 474 trials of training. Figure 4 exhibits the percentage of correct responding and the number of trials applied during training of the first stimuli group for Billy. Simon displayed moderate levels of correct responding throughout the training. Simon did not meet the mastery criterion after 384 trials of training. Figure 5 show the percentage of correct responding and the number of trials during training of the first stimuli group with Simon. Ethan displayed a gradual increase from moderate levels to high levels with each stimuli group. Ethan mastered all of the stimuli groups after a total of 854 trials. Figure 6 exhibits the percentage of correct responding and the number of trials during the training of all stimuli groups with Ethan.

Discussion

In this study the application Superspeak was used to teach receptive labels to five participants with ASD. We examined if it was possible to teach receptive labels through the application and if this generalized to in-vivo objects in a more natural environment. Only one of the participants learned the labels through the tablet and he also showed generalization to other examples of the objects in real life. Four of the participants did not learn any labels and we were not able to test for generalization.

The inability to complete the procedure for these participants can be attributed to several factors. Firstly, these participants were the ones with the lowest level of functioning. This indicates that the procedures in the application were not appropriate for the participants, at the level of independency required by the tablet application. Secondly these participants displayed the highest level of off-task behavior while instructed to work on the tablet. Similar results were found by Chebli et al. (2017), and implies that the amount of off-task behavior is related to the effectiveness of tablet-based interventions. Another possible factor that might have hindered the effectiveness of the procedure is that they demonstrated several attempts to navigate out of the application to access other more desirable applications, such as of games. This could indicate that they were used to navigate freely while using a tablet. In the current study this was not allowed and this may have caused protesting and other off-task behavior. Another factor that may potentially have contributed to the inefficacy of the procedure was the inconsistency of the training. As the school oversaw the daily training sessions, it was sometimes postponed due to illness or vacations from either the participants or the teachers.

Another limitation within the study is that it took the participants longer than planned to complete the procedure, partly due to sickness and vacations, but the main fault most likely lies with the procedure. During training it seemed obvious that there was not enough reinforcement built into the application procedure. The reinforcement strategy heavily depended on generalized reinforcement, in the form of a green check mark and a corresponding sound that was presented after each correct response. Then, after a block of nine correct trials, a piece from a puzzle, blocking a picture, was removed. Once the entire stimuli group was acquired and the procedure completed, all the pieces from the puzzle was removed, and the picture was clearly visible. The problem with this procedure seemed to be that the participants did not get enough contact with the reinforcers for the check mark and correct sound to obtain reinforcing effects in the early stages of the training. This resulted in what seemed to be aimless responding from the participants, as they did not seem to comprehend that the assumed reinforcers indicated correct responding. This could also have been a result from the thorough testing the participants went through before the procedure was implemented. These pre-tests consisted of similar trials as in the training, but without feedback and error-correction. Each participant completed three pre-test sessions, each consisting of nine trials, for each stimuli group. This meant that all participants completed 81 trials on the tablet without receiving any feedback contingent on performance, before the training procedure was implemented. This led to several different response-patterns that might be viewed as superstitious behavior. Some participants tried to press all of the presented pictures at once, some always picked the same spot, for instance always the picture to the right, while some participants always picked the same object independent of which conditional stimulus that was presented. These response-patterns took a while to remove from the participant's behavioral repertoire once the training procedure was implemented and contributed to the initial thin contact with the presumed reinforcers. Pre-testing is however a necessary requirement to ensure good scientific research, even though it sometimes affects the study in a negative manner. Future research should however keep this in mind and keep the feedback-less trials to a minimum. It should be noted that during the study Superspeak updated their application based on the findings from this study, and it now uses a thicker reinforcement schedule where video reinforcers are delivered after three correct responses, in addition to the generalized reinforcers. This update was however not used in this study, in an effort to keep the procedure consistent throughout the training for all the participants. Future researcher should use the updated application to investigate if the updated reinforcement system can improve the rate and quality of learning, as well as the amount of generalization.

Another possible limitation of the procedure is the incorporated error-correction procedure. One of the problems with this procedure is that the reinforcement was the same for prompted and correct trials. This is specially an issue as it could lead to prompt dependency and confusion, and hinder the participants from advancing in the procedure. Another limitation with the error-correction procedure may be that the stimulus positions did not change under the prompted trials. This can lead to faulty stimulus control where the controlling variable becomes the position of the stimulus, instead of the conditional stimulus. It is also not common to have three stages of prompted trials before the participants have the opportunity to respond independently, as this could also lead to prompt dependency. Future research should therefore implement changes to these aspects of the error-correction procedure and adopt a procedure more commonly seen in discrete trail training to examine if it leads to an improved rate of acquisition.

The results from this study also indicate that tablet-based interventions can be taught on a tablet and that it may lead to improved skills in the natural environment as well as in the artificial tablet setting. The participant that completed the study displayed relative quick acquisition of the receptive language skills and showed high rates of generalization to in-vivo objects for the first two stimuli groups, as well as high levels of correct responding during maintenance testing. This participant did however show lower levels of generalization in the last stimuli group, despite the fact that he demonstrated a quick and thorough acquisition in the tablet condition. The fact that some objects, in this case clothes, can be harder to generalize from the tablet training to in-vivo objects is interesting, and should be targeted in future research. This can for instance be accomplished by applying the same stimuli groups as the ones utilized in this project, with different participants within the same range of functioning, and see if similar results are obtained. Future research could also vary the level of functioning of the participants, to see if there is a relation between the level of functioning and the degree of generalization to different stimuli groups. This was one of the targets for this study but could not be accomplished due to the fact that only one student could be tested for generalization.

The effectiveness of the intervention for Ethan might be attributed to several factors. Ethan already had experience from operating tablets. This acquaintance with tablets could shorten the adjustment period to the tablet, and increase the likelihood of teacher independence. However the tablet related experience came from playing various games in his leisure time, and not on-task work and could therefore also be regarded as a disadvantage due to the application restrictions and the novel nature of the application. This specific participant also had some experience with independent assignments, where the teacher would start a task and withdraw from the situation until the task was completed, then deliver reinforcement. The participant did not, however, have any significant experience with fully independent assignments, like the ones employed in this study.

The results from this study also indicate that, for some children, tablets can indeed promote independence. The participant that completed the full training procedure worked mostly without teacher assistance. The teacher would simply start the application, and then let the participant work for the pre-determined duration before ending the session by pausing the application and removing the tablet. These results are consistent with Chebli et al. (2017), and even though this does not necessarily mean that it would be possible to remove teachers from the training, it indicates that, at least, one teacher can, in some cases, supervise multiple students. It also suggests that some of the teaching can be done through tablets without the presence of a teacher or a therapist, for instance at home with the supervision from a caregiver. However, the participants that did not finish the procedure showed less independence. They had more attempts at elopement and off-task behavior. This might indicate that the degree of independent tablet training is related to the level of functioning of the children, and may suggest that for some children tablet training will require a higher level of teacher presence at the beginning, before being gradually faded later in the training program. However, because there only was one participant that completed the entire procedure, these findings are hard to generalize to other children, and should be investigated further.

Another limitation with this study is the fact that four of the participants did not complete the full procedure and was therefore not able to be tested for generalization. This led to a reduction in the external validity of the study and limited the generalization value of these results to other children. Two of the other participants did not show any improvement in labeling after 474 and 384 trials. The intervention was therefore stopped, as it was deemed unethical to continue. The other two participants that did not finish the teaching were not able to complete the training within the allotted time period. They changed schools and were not available to further participation.

This study indicates that for some, but not all children with ASD, tablet-based intervention can be effective in training receptive language and that it may yield generalization to natural settings. This study also shows that tablet training is not always successful, and that there is still much research necessary to develop a thorough understanding of the variables controlling its effectiveness. The study demonstrates the necessity of incorporating the knowledge from the literature and experienced behavior analysts, to tailor prompting-, error correction procedures and schedules of reinforcement to ensure effective training procedures.

References

Achmadi, D., Kagohara, D. M., van der Meer, L., O'Reilly, M. F., Lancioni, G. E., Sutherland, D., . . . Sigafoos, J. (2012). Teaching advanced operation of an iPod-based speech-generating device to two students with autism spectrum disorders. *Research in Autism Spectrum Disorders, 6*(4), 1258-1264.

AmericanPsychiatricAssociation. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

AutismSpeaks. (2017). *Autism Apps.* Retrieved from https://www.autismspeaks.org/autism-apps?page=5&tid_1=39876&tid=All&tid_2=All&keys=

Bhat, S., Acharya, U. R., Adeli, H., Bairy, G. M., & Adeli, A. (2014). Autism: cause factors, early diagnosis and therapies. *Reviews in the Neurosciences, 25*(6), 841-850.

Bosseler, A., & Massaro, D. W. (2003). 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), 653-672.

Burton, C. E., Anderson, D. H., Prater, M. A., & Dyches, T. T. (2013). Video self-modeling on an iPad to teach functional math skills to adolescents with autism and intellectual disability. *Focus on Autism and Other Developmental Disabilities, 28*(2), 67-77.

Chebli, S. S., Lanovaz, M. J., & Dufour, M.-M. (2017). Generalization Following Tablet-Based Instruction in Children With Autism Spectrum Disorders. *Journal of Special Education Technology, 32*(2), 70-79.

Clarke, B., & Svanaes, S. (2014). An updated literature review on the use of tablets in education. *Tablets for Schools. UK: Family Kids & Youth*.

DiGennaro Reed, F., D. , Hyman, S., R. , & Hirst, J., M. (2011). Applications of Technology to Teach Soscial Skills to Children with Autism. *Research in Autism Spectrum Disorders, 5*, 1003-1010.

Egeberg, G., Hultin, H., & Berge, O. (2017). *Monitor skole 2016: skolens digitale tilstand* (2 ed.): Senter for IKT i utdanningen.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2010). Using participant data to extend the evidence base for intensive behavioral intervention for children with autism. *American Journal on Intellectual and Developmental Disabilities, 115*(5), 381-405.

Flores, M., Musgrove, K., Renner, S., Hinton, V., Strozier, S., Franklin, S., & Hil, D. (2012). A comparison of communication using the Apple iPad and a picture-based system. *Augmentative and Alternative Communication, 28*(2), 74-84.

Ganz, J. B., Boles, M. B., Goodwyn, F. D., & Flores, M. M. (2014). Efficacy of handheld electronic visual supports to enhance vocabulary in children with ASD. *Focus on Autism and Other Developmental Disabilities, 29*(1), 3-12.

Jowett, E., Moore, D., & Anderson, A. (2012). Using an iPad-based video modelling package to teach numeracy skills to a child with an autism spectrum disorder. *Developmental neurorehabilitation, 15*(4), 304-312.

Kagohara, D. M., Sigafoos, J., Achmadi, D., O’Reilly, M., & Lancioni, G. (2012). Teaching children with autism spectrum disorders to check the spelling of words. *Research in Autism Spectrum Disorders, 6*(1), 304-310.

Kagohara, D. M., van der Meer, L., Achmadi, D., Green, V. A., O'Reilly, M. F., Mulloy, A., . . . Sigafoos, J. (2010). Behavioral intervention promotes successful use of an iPod-based communication device by an adolescent with autism. *Clinical Case Studies, 9*(5), 328-338.

Klintwall, L., Eldevik, Sigmund., Eikeseth, Svein. (2015). Narrowing the gap: Effects og intervention on developmental trajectories in autism. *Journal of Autism and Development Disorders, 19*, 53-63.

Knight, V., McKissick, B., R. , & Saunders, A. (2013). A Review of Technology-Based Interventions to Teach Academic Skills to Students with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 43*, 2628-2648. doi:10.1007/s10803-013-1814-y

Lovaas, O. I. (2003). *Teaching Individuals With Developmental Delays: Basic Intervention Techniques*. 8700 Shoal Creek Blvd., Austin, TX: Pro-ed Inc.

Murdock, L. C., Ganz, J., & Crittendon, J. (2013). Use of an iPad play story to increase play dialogue of preschoolers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*(9), 2174-2189.

Osnes, P. G., & Lieblein, T. (2003). An Explicit Technology of Generalization *The Behavior Analyst, 3*(4), 364-374.

Sansosti, F. J., & Powell-Smith, K. A. (2008). Using computer-presented social stories and video models to increase the social communication skills of children with high-functioning autism spectrum disorders. *Journal of Positive Behavior Interventions, 10*(3), 162-178.

Schery, T., & O'Connor, L. (1997). Language intervention: Computer training for young children with special needs. *British Journal of Educational Technology, 28*(4), 271-279.

Shah, N. (2011). *Ipads Become Learning Tools for Students with Disabilities.* Retrieved from https://www.edweek.org/ew/articles/2011/03/02/22ipad.h30.html

Still, K., May, R. J., Rehfeldt, R. A., Whelan, R., & Dymond, S. (2015). Facilitating derived requesting skills with a touchscreen tablet computer for children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 19*, 44-58.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of applied behavior analysis, 10*(2), 349.

van der Meer, L., Kagohara, D., Achmadi, D., Green, V. A., Herrington, C., Sigafoos, J., . . . Rispoli, M. (2011). Teaching functional use of an iPod-based speech-generating device to individuals with developmental disabilities. *Journal of Special Education Technology, 26*(3), 1-11.

Van Laarhoven, T., Johnson, J. W., Van Laarhoven-Myers, T., Grider, K. L., & Grider, K. M. (2009). The effectiveness of using a video iPod as a prompting device in employment settings. *Journal of Behavioral Education, 18*(2), 119-141.

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical psychology review, 30*(4), 387-399.

Özerk, K. (2016). The Issue of Prevalence of Autism/ASD. *International Electronic Journal of Elementary Education, 9*(2), 263-306.

Tables

Table 1

The tablet-based training procedure

Step	Procedure
Step 1	Sounds --> Picture set 1
Step 2	Test of relation 1
Step 3	Sounds --> Picture set 2
Step 4	Test relation 2
Step 5	Mix relation 1 og 2
Step 6	Test relation 1 and 2

Table 2

Error-correction procedure

Step	Procedure
Step 1	Fading in color for the wrong options Brightening in color for the correct picture Model for choosing the right option.
Step 2	Fading in color for the wrong options Nudge of the correct picture
Step 3	Enlargement of the correct option
Step 4	Regular trial

Figures

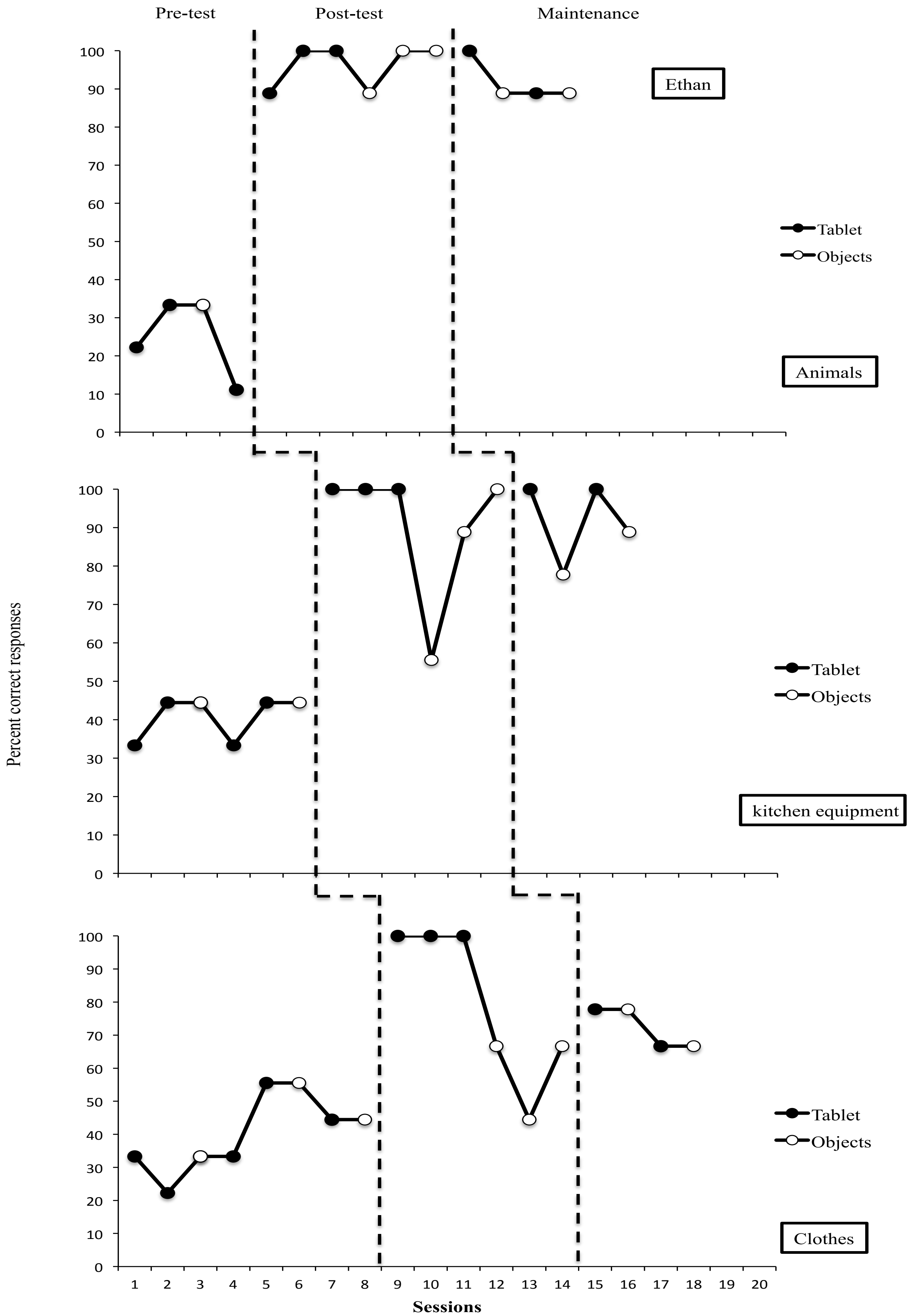

Figure 1. The percent of correct responses for Ethan during pre-test, post-test, and maintenance conditions for the three stimuli groups. The y-axis displays the percent of correct responses while the x-axis shows the number of sessions.

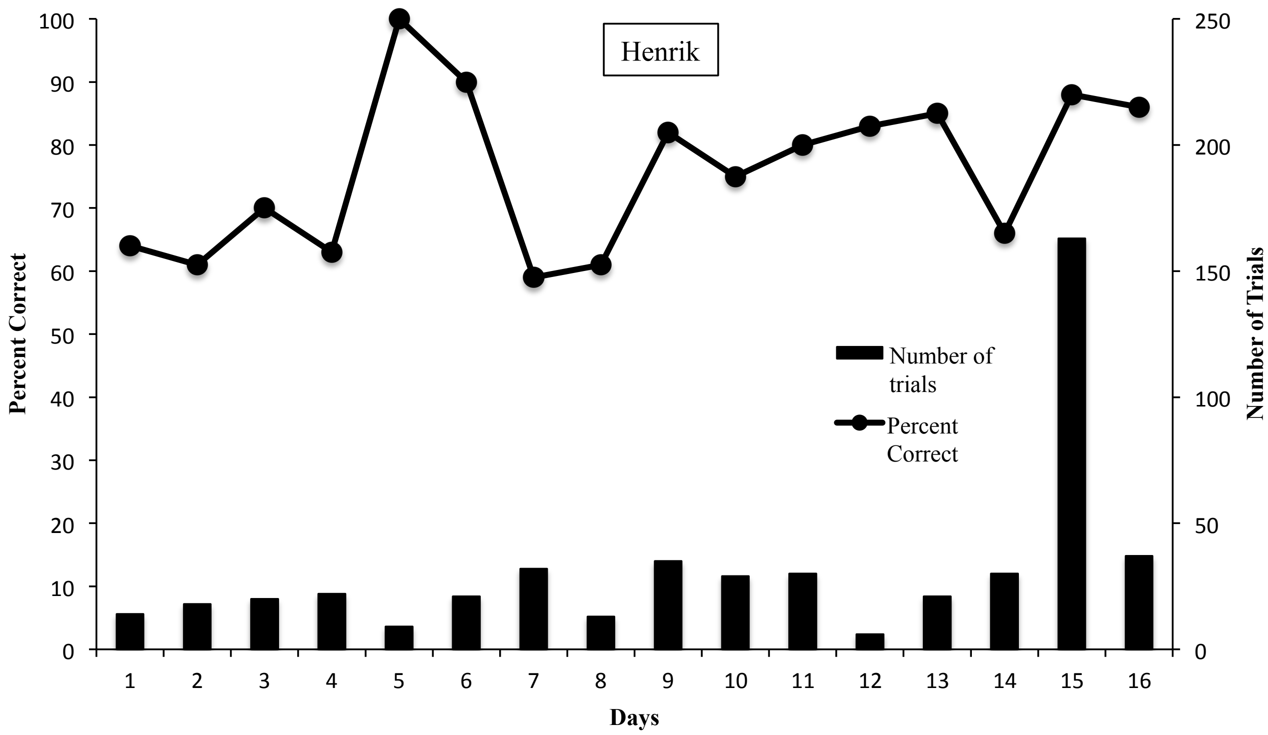

Figure 2. The number of trials and the percent of correct responding during training of the first stimuli group with Henrik. The first y-axis displays the percent of correct responding and the second y-axis shows the number of trials. The x-axis displays the day of the training.

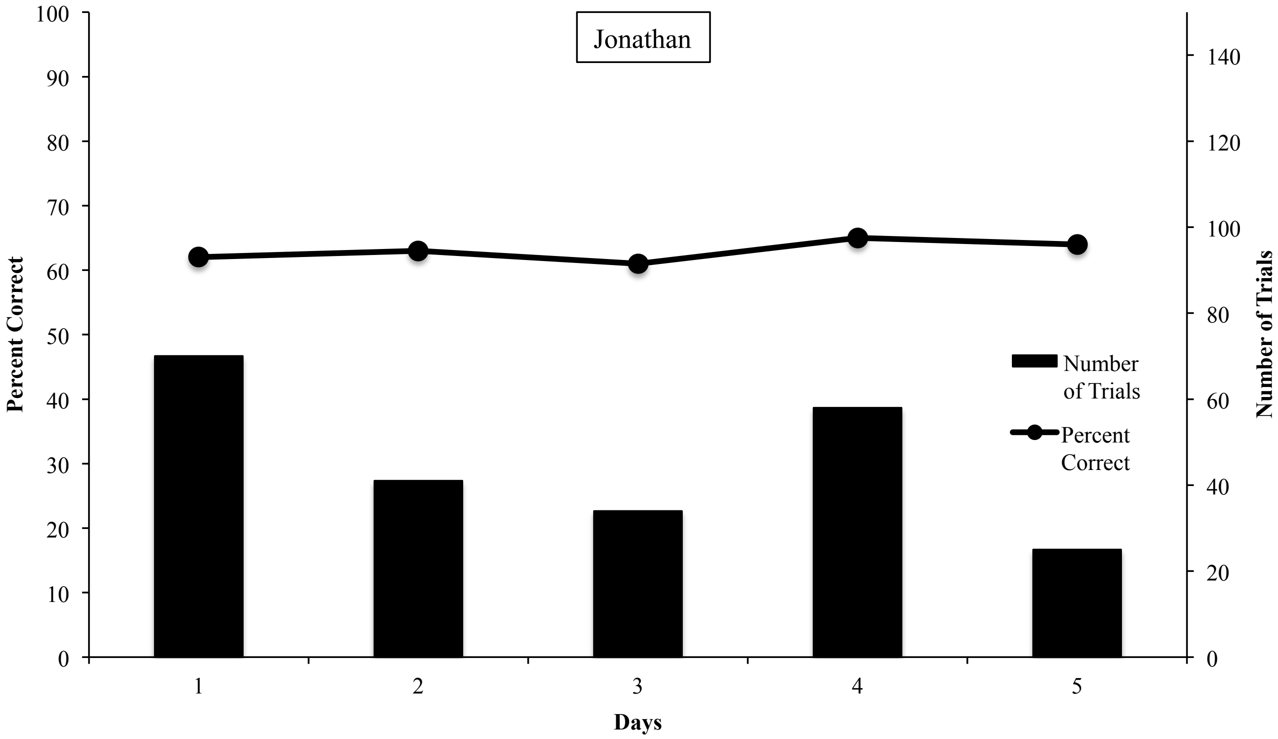

Figure 3. The number of trials and the percent of correct responding during training of the first stimuli group with Jonathan. The first y-axis displays the percent of correct responding and the second y-axis shows the number of trials. The x-axis displays the day of the training.

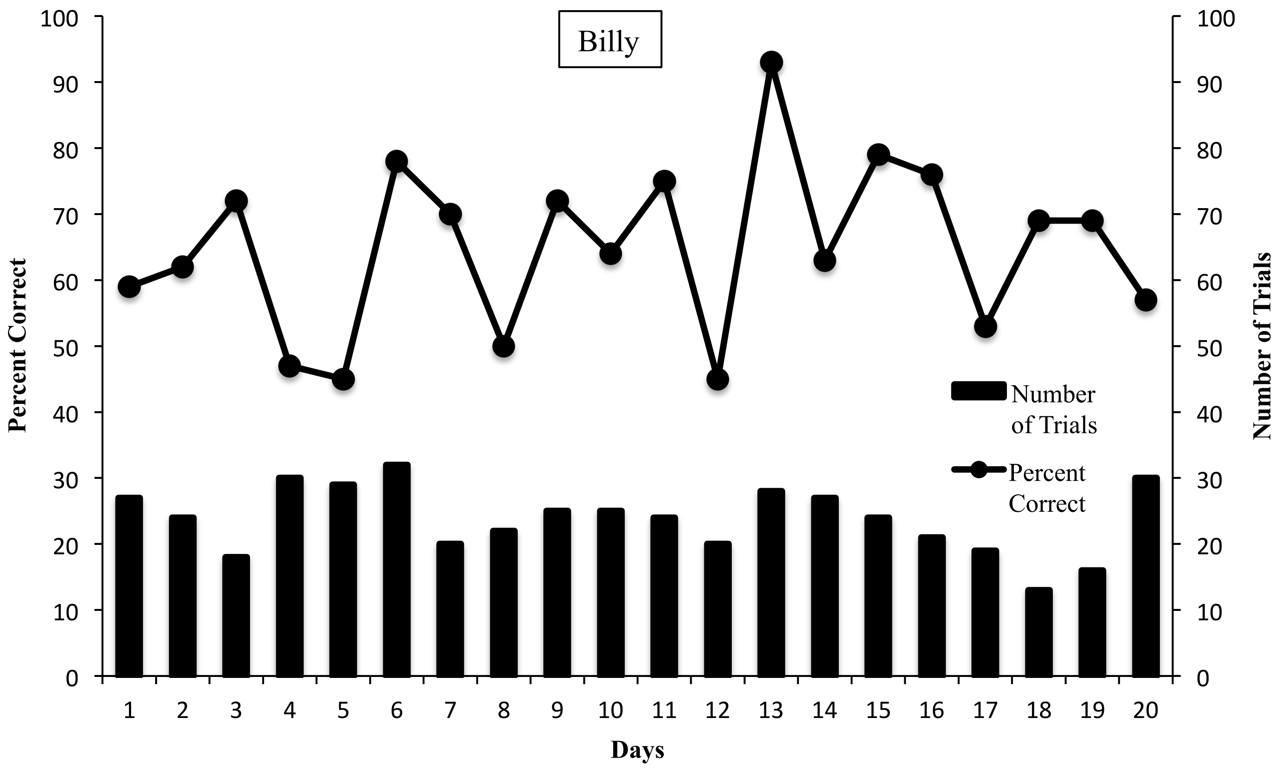

Figure 4. The number of trials and the percent of correct responding during training of the first stimuli group with Billy. The first y-axis displays the percent of correct responding and the second y-axis shows the number of trials. The x-axis displays the day of the training.

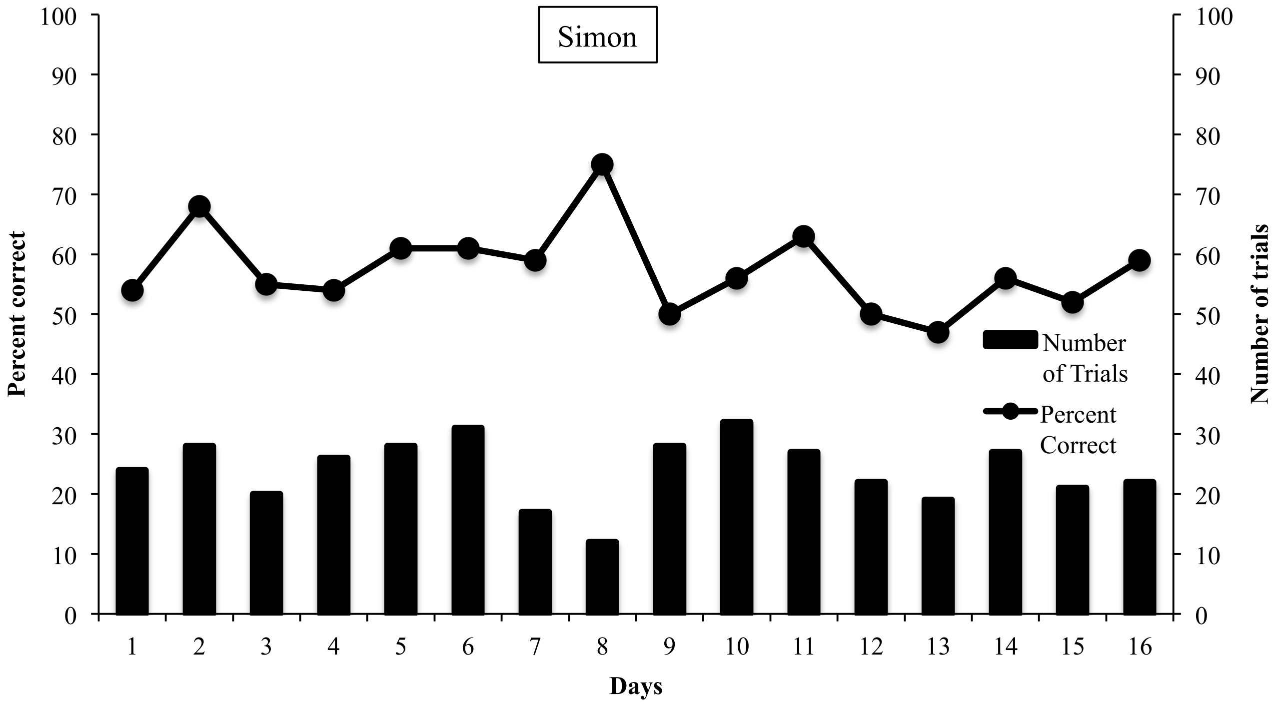

Figure 5. The number of trials and the percent of correct responding during training of the first stimuli group with Simon. The first y-axis displays the percent of correct responding and the second y-axis shows the number of trials. The x-axis displays the day of the training.

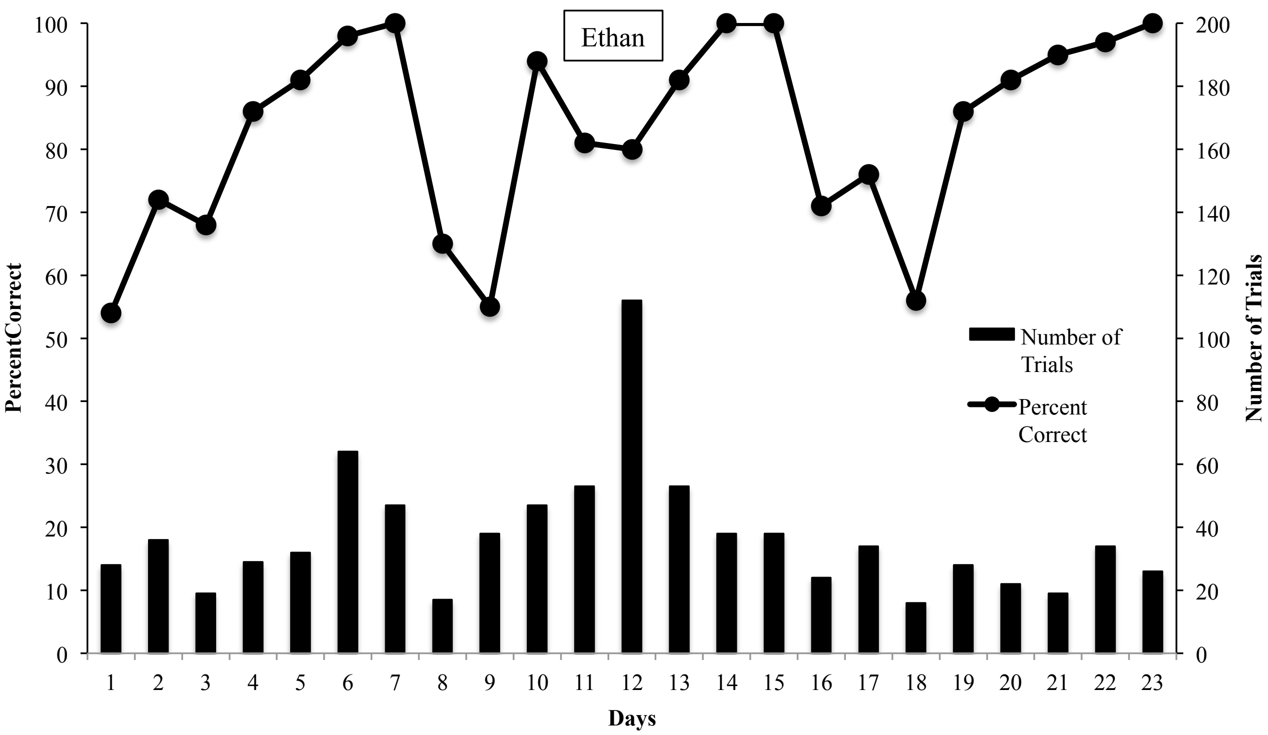

Figure 6. The number of trials and the percent of correct responding during training of all stimuli groups with Ethan. The first y-axis displays the percent of correct responding and the second y-axis shows the number of trials. The x-axis displays the day of the training.

Appendix

Forespørsel om deltakelse i forskningsprosjektet

" Overføring av språk lært på en Ipad til det virkelige liv."

Bakgrunn og formål

Formålet med dette studie er å måle overføring av språk etablert gjennom tablet applikasjonen Superspeak til det virkelige liv, hos barn med autisme. Dette gjennomføres for å undersøke om tablet applikasjoner faktisk fører til mer funksjonelle ferdigheter i et naturlig miljø. Dette er essensielt ettersom det stadig kommer flere og flere applikasjoner av denne typen på markedet og elver ofte bruker mye opplæringstid på dette. De potensielle fordelene ved slike applikasjoner er vesentlige, ettersom de kan tillate store deler av treningen til å forekomme uten en terapeut er tilstede, noe som
drastisk kan redusere de samlede ressursene som er nødvendig for å gi god opplæring. Dette prosjektet er en del av en masteroppgave ved fakultetet for adferdsvitenskap på høgskolen i Oslo og Akershus

Du forespørres om å delta i studien ettersom læreren til barnet ditt vurderer at dette prosjektet vil være til nytte for barnet. Dette kommer av at ditt barn allerede mottar lignende opplæring, og dette prosjektet kan bidra til å effektivisere dette.

Hva innebærer deltakelse i studien?

Deltagelsen i studien innebærer at barnet gjennomfører språktrening gjennom Ipad applikasjonen Superspeak. Denne treningen går ut på at barnet blir instruert om å trykke på et spesifikt objekt, og blir presentert et valg mellom tre forskjellige bilder. Treningen inkluderer en rekke bilder og korte videosnutter som benyttes for å gjøre treningen underholdene for barnet. Alle objekter som benyttes i studien vil bestemmes i samarbeid med barnets lærer, slik at det samsvarer med barnets aktuelle læringsmål. Treningen vil bestå av to økter på omtrent 10 minutter daglig, og vil vare mellom 3-10 uker avhengig av hvor raskt barnet når målet for mestring. Studien kan også inneholde en kartlegging av barnets IQ skåre, en Assessment of Basic Language and Learning skills-Revised (ABLLS-R) samt en Childhood Autism Rating Scaleen (CARS) kartlegging dersom disse ikke allerede foreligger. Foreldre kan på forespørsel få tilgang til den aktuelle applikasjonen som benyttes i treningen.

Hva skjer med informasjonen om deg?

Alle personopplysninger vil bli behandlet konfidensielt. De eneste som vil ha tilgang til personopplysninger vil være prosjektgruppen, bestående av en veileder og en student. Alle personopplysninger vil være sikret med et brukernavn og passord for å ivareta konfidensialitet.

Deltagere vil kun være gjenkjennelig i publikasjon indirekte ved deres aktuelle testskårer. Ingen navn eller andre personopplysninger vil bli benyttet i publikasjon.

Prosjektet skal etter planen avsluttes 17.12.2017. Alle personopplysninger og datamaterialet anonymiseres ved prosjektslutt.

Frivillig deltakelse

Det er frivillig å delta i studien, og du kan når som helst trekke ditt samtykke uten å oppgi noen grunn. Dersom du trekker deg, vil alle opplysninger om deg bli anonymisert. Dersom du ikke ønsker å delta i studien, eller bestemmer deg for å trekke deg ved en senere anledning vil ikke dette påvirke ditt forhold til lærere eller andre personer ved din aktuelle utdanningsinstitusjon.

Dersom du ønsker å delta eller har spørsmål til studien, ta kontakt med Vetle Berge ved mail: vberge@hioa.no eller ved telefon: 41603223. Veileder Sigmund Eldevik kan nås på mail: sigmund.eldevik@hioa.no.

Studien er meldt til Personvernombudet for forskning, NSD - Norsk senter for forskningsdata AS.

Samtykke til deltakelse i studien

Jeg har mottatt informasjon om studien, og er villig til å delta

----
(Signert av prosjektdeltaker, dato)