

Case Study 1: An Evidence-Based Practice Review Report

Theme: School Based Interventions for Learning

How effective is the 'Lego® Therapy' intervention to support children aged 5-16 with an autistic spectrum disorder with their social communication difficulties in school?

Summary

LEGO® Therapy is an intervention that has started to emerge in schools to help children with Autistic Spectrum Disorder (ASD) in their social communication difficulties. Originally proposed by LeGoff (2004) this intervention employs the use of three key roles 'Engineer' 'Builder' and 'Supplier' to build a model together using LEGO® bricks. This is facilitated by a trained adult, although child led, to allow the children a means of developing their social communication skills. The aim of the review is to determine how effective the intervention is at helping children with ASD in their social communication skills. A systematic literature search identified five studies which met the inclusion criteria. These were coded against an adapted protocol form (Kratchowill, 2003) to evaluate their weight of evidence (Gough, 2007). Four out of five studies achieved an Overall Weight of Evidence score of Medium, however, there were concerns on the underpowered nature of some of the studies and gender biases. Large effect sizes were displayed, although some studies were underpowered. Recommendations on how Educational Psychologists could use the findings of the review were discussed.

Introduction

Lego® Therapy

Exploiting a child's natural curiosity and motivation within LEGO® to dictate a behavioural change, is the fundamental ethos of Lego® Therapy (Owens, Granader, Humphrey and Baron-Cohen, 2008). It is a child-led and peer-based intervention that builds upon the child's interests in construction play to promote a willingness to collaborate and interact (LeGoff, Krauss and Levin, 2010). Lego® Therapy was originally developed by US psychologist Daniel LeGoff (2004) after observing two of his ASD clients, who had shown poor motivation to interact, were playing with Lego® in his waiting area and displaying positive interactions. LeGoff (2004) then devised a role based building intervention through the medium of Lego®. This resulted in an increase in interaction, which in turn promoted the acquisition of key skills such as sharing, collaboration, conflict resolution, verbal and non-verbal communication (Owens et al., 2008). The original intervention was solely aimed at children with ASD, however, it has also been found to benefit many children with communication and social developmental difficulties (LeGoff, 2004).

Lego® Therapy is delivered in thirty-minute sessions, once per week, in which a triage of children jointly build Lego® models. The group first develop a name for the club, usually 'Lego® Club', alongside general session rules. The children take it in turns to fulfil specific roles:

- Engineer – Uses plans to describe the construction needed and bricks.

- Supplier – Finds the correct bricks as described by the Engineer and gives to Builder.
- Builder – Uses the verbal instructions given by the Engineer to put the bricks together.

The triage must use their communication and social skills to build the model. The session is facilitated by an adult, usually Teaching assistant (TA), who supports the group in developing language and taking ownership of the build (LeGoff, 2004). The activity provides opportunities for children to use their joint problem solving, joint creativity, joint attention, verbal and non-verbal communication skills (LeGoff, Krauss and Levin, 2010). External rewards are not needed as participation in the group is rewarding enough (LeGoff, 2004).

Psychological Basis

Social Communication

Social communication is *“the intersection of language and social behaviours observed during peer interactions... that is, the verbal and nonverbal behaviours children display as they approach peers, maintain conversations, and resolve conflicts during peer interactions”* (Timler, Olswang, and Coggins, 2005, p. 171). It has been proposed that play can facilitate a child to practice their interaction skills which in turn develops their social communication abilities (Walberg and Craig-Unkefer, 2010).

ASD and Social Communication

Autistic Spectrum Disorder (ASD) is found to be a neurodevelopmental disorder (Hart, 2011). One must meet a set criteria for a diagnosis of ASD (DSM-V, American

Psychiatric Association (APA), 2013). Limitations within communication and social-interaction are two of the many criteria used for a diagnosis of ASD (APA, 2013). It has been found that children with ASD have a weakness in understanding the emotional state of others and comprehending verbal and non-verbal communication, this can lead to difficulties in social communication (National Research Council, 2001). The lack of social communication has then been linked to difficulties in being included in collaborative play and forming relationships (National Research Council, 2001). This is supported by research which found that children with autism, who used very little language, had a greater difficulty in play and socialisation than the children with autism who had a better use of language (Lord and Pickles, 1996).

Lego® Therapy

It has been found that children with ASD are highly motivated to participate in Lego® Therapy with no external rewards (LeGoff 2004). Owens et al. (2008) suggests that children with ASD may be attracted to Lego due to Baron-Cohen's hyper-systemizing theory (Baron-Cohen, 2006). It has been suggested that children with ASD have a strong urge to systemize - to predict patterns and changes in lawful events (Baron-Cohen, 2008). Baron-Cohen (2006) suggested that the systemizing mechanism within ASD children allow them to look for input-operation-output relationships, whilst being able to derive laws and patterns from these relationships. Consequently, playing with Lego® appeals to children with ASD as the toy itself is suited to being systemized due to its predictable and systematic nature (Owens et al., 2008). Facilitated playing can have a significant impact on a multitude of skills all linked with social communication (Walberg and Craig-Unkefer, 2010). These include the development of social

interactions and awareness to make confusing situations or interactions easier to understand (Bodrova and Leong, 2003; Goncu, 1999).

Rationale and Relevance to Educational Psychology Practice

Inclusion into mainstream has been a key focus of governmental initiatives which include children with ASD (DoE, 2001, p. 360). It has been suggested that children with ASD who are fully included within mainstream, tend to have better social interaction and engagement than those in segregated environments (Harrower and Dunlap, 2001). However, it has been expressed by educators that they often find it challenging to include their ASD students as it is beyond their capabilities (Spears, Tollefson, and Simpson, 2001). The shift in school focus from solely academic achievement to greater social and emotional development has led to a climate of joint approaches from professionals (including Educational Psychologists) and schools (Children and Families Act, 2014; DfES, 2005). This was derived primarily from the work of Goleman (1996) who suggested that social and emotional abilities could play a greater influence on personal and academic success.

Interventions are used to support children with emotional and social competences which would in turn promote resiliency for the future (Stanton-Chapman, Denning and Jamison, 2012). There have been several initiatives, strategies and interventions which aim to promote inclusion (McConnell, Missall, Silberglitt and McEvoy, 2002). However, it has been suggested that particularly for children with ASD the evidence bases for these interventions are generally weak (Howlin, Gordon, Pasco, Wade. and Charman, 2007).

Within this climate of an increased focus on social and emotional development in UK

schools, it is increasingly common for a school to seek advice and interventions from Educational Psychologists (Children and Families Act, 2014; DfES, 2005). It would be appropriate for an Educational Psychologist to recommend, train and support staff with interventions such as LEGO® Therapy to address these concerns. It has been observed that building social competence and communicational skills is a key focus in schools (Stanton-Chapman et al., 2012). A strength of LEGO® therapy is that the Educational Psychologist can train the school in this intervention which in turn builds capacity for the school to address its social, emotional and communicational obligations. Furthermore, the interventions that the Educational Psychologist recommends are increasingly being scrutinised for their evidence. This has been linked to the shift to evidence based practice in the field of Educational Psychology (Fredrickson, 2002). It would therefore be beneficial both academically and professionally, to thoroughly explore the evidence-base of interventions such as LEGO® Therapy. Through this systematic literature review of LEGO® therapy, the claims of supporting children with ASD with their social communication difficulties (LeGoff, 2004) can be effectively evaluated.

Review Question

How effective is the 'Lego® Therapy' intervention to support children aged 5-16 years with an autistic spectrum disorder with their social communication difficulties in school?

Section 3: Critical Review of the Evidence Base

1. Literature search

An in-depth literature search was carried out to address the research question. An initial search was performed on 28th January 2017 using the online databases PsychINFO, ASSIA, ERIC and Wiley. *Table 1* shows the search terms used for each database.

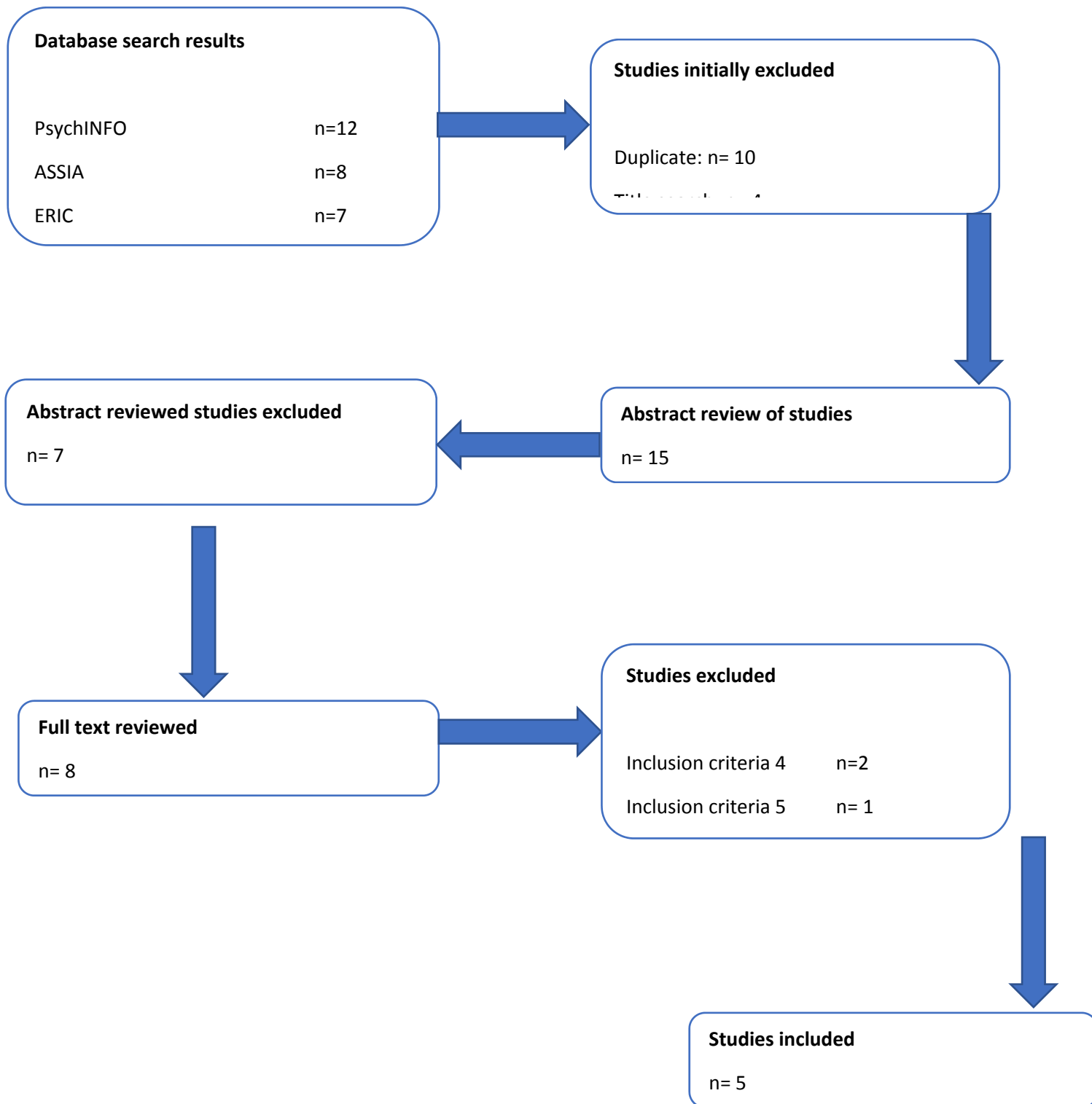
Table 1

Database and search terms

Data Base	Search Terms
<i>PsychINFO</i>	'lego therapy'
<i>ASSIA</i>	'lego therapy'
<i>ERIC</i>	'lego therapy'
<i>Wiley</i>	'lego therapy' (Abstracts only)

Figure 1 displays a flowchart which offers more information pertinent to the literature search. The inclusion and exclusion criteria for the searches are defined in *Table 2*. *Table 3* lists the full names of finalised studies used in this review. Those studies that were excluded after a full text review can be found in *Appendix 1*. A summary of all the studies used in this review can be found in *Appendix 2*.

Figure 1: Literature Search Flow Diagram



2. Table 2

3. Inclusion and Exclusion criteria

		Inclusion Criteria	Exclusion Criteria	Justification
1	Participants	Participants must be of school age (5 -16) and have a diagnosis of Autistic Spectrum Disorder (ASD)	Participants are not between the ages of 5 -16 or they do not have a diagnosis of ASD	To ensure that all participants have a clinically recognised diagnosis of ASD
2	Type of Publication	Studies must be in a published peer reviewed journal.	Not a published peer reviewed study.	To ensure that the quality of studies used is of a high calibre and has been subjected to scrutiny
3	Language	Studies are written in English	Studies not written in English	To ensure research is understood
4	Intervention	Intervention must be LEGO® Therapy as described by LeGoff (2004) including roles as ‘engineer’ ‘builder’ and ‘supplier’	LEGO® based interventions that are not ‘LEGO therapy’ and do not follow LeGoff (2004) model.	To ensure that the intervention reviewed is the LEGO® Therapy intervention as described by LeGoff (2004)
5	Measures	Study must have pre and post intervention data	Study lacks pre and post data	To ensure that a measurement of the effectiveness can be established
6	Type of Design	Must be a group based or single case study design	Not a group based or single case design	To ensure a comparison of between or within participants can be ascertained

4. Table 3

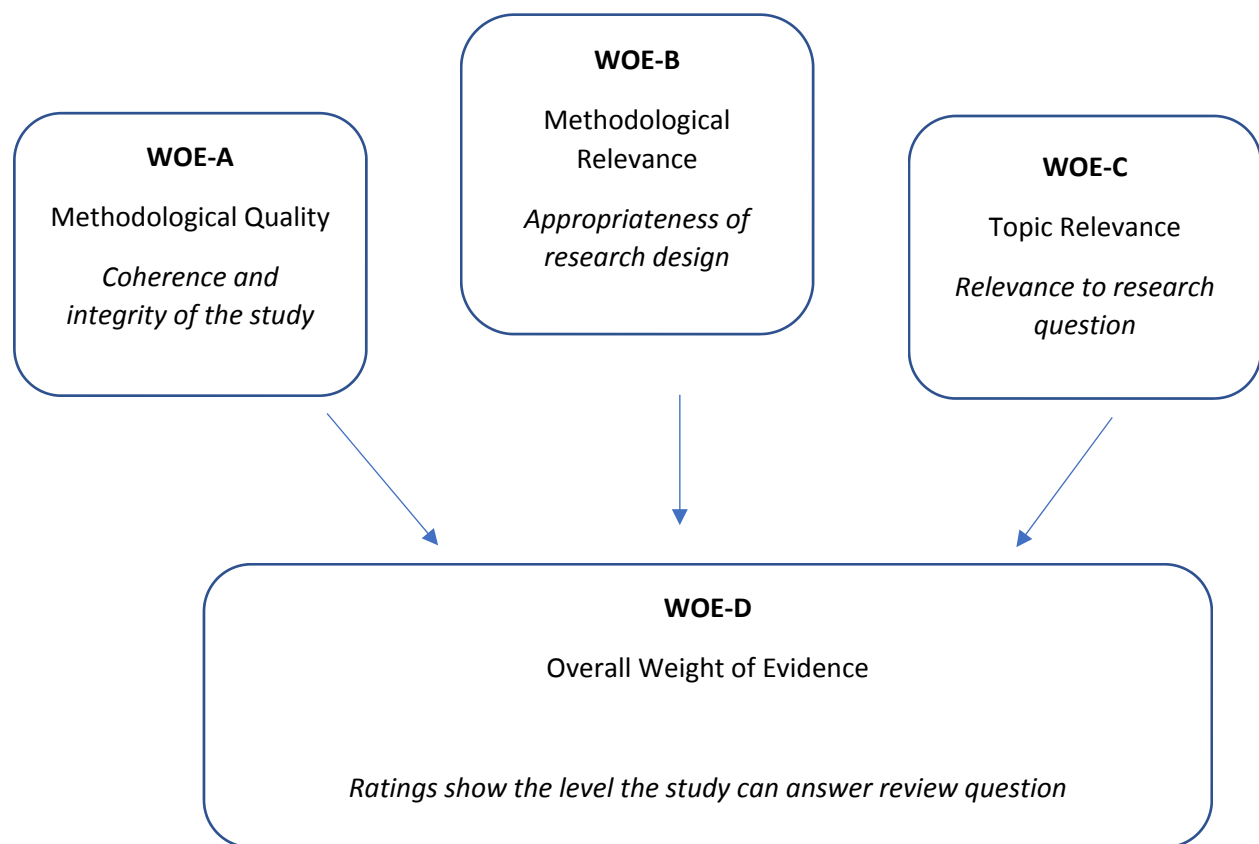
5. *Selected Studies in Review*

Included Studies	
1	Barakova, E. I., Bajracharya, P., Willemsen, M., Lourens, T., and Huskens, B. (2014). Long-term LEGO therapy with humanoid robot for children with ASD. <i>Expert Systems</i> , 32(6), 698–709.
2	Owens, G. Granader, Y. Humphrey, A. and Baron-Cohen, S. (2008) LEGO Therapy and the Social Use of Language Programme: An Evaluation of Two Social Skills Interventions for Children with High Functioning Autism and Asperger Syndrome, <i>Journal of Autism Developmental Disorders</i> , 38, 1944-1957.
3	LeGoff, D. and Sherman, M. (2006) Long-term outcome of social skills intervention based on interactive LEGO play, <i>Autism</i> , 10, (4), 317-329.
4	LeGoff, D. (2004) Use of LEGO as a Therapeutic Medium for Improving Social Competence, <i>Journal of Autism and Developmental Disorders</i> , 34, (5), 557-571.
5	Andras, M. (2012) The value of Lego® Therapy in promoting social interaction in primary aged children with autism. <i>Good Autism Practice</i> , 13, (2), 18 – 25.

Critical Appraisal of Studies

Gough’s (2007) Weight of Evidence Framework was used to critically appraise the five studies selected. This allowed for a detailed systematic appraisal of the quality and relevance of the studies, whilst maintaining an objective stance. The five studies were rated within three domains; methodological quality (Weight of Evidence A), methodological relevance (Weight of Evidence B) and topic relevance (Weight of Evidence C). These scores would then be used to determine a final score for the evidence (Weight of Evidence D). *Figure 2* illustrates this process in greater detail.

6. Figure 2: Gough’s (2007) Weight of Evidence Framework.



7. Methodological Quality (WOE-A)

Methodological quality was determined by coding the studies against a coding protocol. The UCL Educational Psychology Literature Review Coding Protocol, adapted from the APA Task Force Coding Protocol by Kratochwill (2003) was used for group designs. Single case studies were coded against the Quality Indicators: Within Single-Subject Research by Horner et al. (2005) and some scales from Kratochwill (2003). After completion of coding the studies were given a score between 1 and 3 which was assigned to its methodological weight: Low (1.4 or below), Medium (1.5 to 2.5) or High (2.6 to 3). For further information see *Appendix 3*.

8. Methodological Relevance (WOE-B)

To determine Methodological Relevance two different criteria were used for weighting single case and group comparison (see *Appendix 3*). For the single case criteria, a greater focus on multiple baselines was sought as this would allow a greater differentiation on the effectiveness of the intervention (Horner et al., 2005). Within the group comparison criteria, a greater emphasis was put on randomised controlled trials. This design was pursued as it would minimise researcher and selection bias (Guyatt et al., 1995). Studies were given a score of 1 (low weighting), 2 (medium weighting) and 3 (high weighting) with regards to their methodological relevance.

This Literature review will not discriminate against the methodological differences between single case study and comparison groups. There have been numerous criteria developed to measure effectiveness within the field of education and psychology, some of these include single-subject design (Chambless et al., 1998; Horner et al., 2005). Although traditionally the use of Randomised Control Trials (RCT) is described as the gold standard of psychological research, there is a strong case to be had for single case designs (Horner et al., 2005). Kazdin (2010) suggests that ideally there should be several studies which employ a variety of methodological stances to establish if an intervention is evidenced based. Exclusively using RCTs to establish the evidence base of an intervention can cause misinterpretations that it is effective with most or all participants, and ignore that there could be responders and non-responders (Kazdin, 2010). Therefore, systematic evaluation at an individual level is needed, especially in clinical or educational practices (Kazdin, 2010). Kazdin (2010) suggests that the

baseline in single case designs establishes a reputable benchmark from which subsequent conditions can be compared for effectiveness. As this review is for both research and practice purposes, on how LEGO® therapy is affecting children within a school setting, then both methodological stances have been deemed appropriate in expressing causality. Therefore, the individually created WOE B criteria will assess the two methods on separate criteria in relation to answering the research question.

9. Topic Relevance (WOE-C)

The score for topic relevance was ascertained by focusing on the following three key areas: sample type, methods of assessment and intervention delivery. These findings relate strongly to the research question of this review. The weightings for WOE-C reflect how strongly the studies have been judged to answer the research question. Studies were given a score of 1 (low weighting), 2 (medium weighting) and 3 (high weighting). See *appendix 3* for a full description of the criteria used.

10. Weight of Evidence Findings

The overall weight of evidence (WOE-D) was determined by the average of the WOE A, B and C scores. This score can then be used to determine to what extent the study can answer the research question. WOE is separated into 3 scoring ranges Low (1.4 or below), Medium (1.5 to 2.5) or High (2.6 to 3) overall weights. *Table 4* presents the WOE findings for the studies used in this review.

11. Table 4

12. Weight of Evidence Findings

Study ID	Authors	WOE – A (Methodological Quality)	WOE – B (Methodological Relevance)	WOE – C (Topic Relevance)	WOE – D (Overall)
1	Barakova et al. (2014)	1.3 Low	1 Low	1 Low	1.1 Low
2	Owens et al. (2008)	1.6 Medium	3 High	2 Medium	2.2 Medium
3	LeGoff and Sherman (2006)	1.6 Medium	2 Medium	2 Medium	1.9 Medium
4	LeGoff (2004)	1.3 Low	2 Medium	2 Medium	1.8 Medium
5	Andras (2012)	1.3 Low	1 Low	3 High	1.8 Medium

Critical Review

13. Participants

Researchers Interests

It must be noted that two of the five studies used in this review were implemented by the creator of the LEGO® therapy intervention (LeGoff, 2004; LeGoff and Sherman, 2006). Within these studies, it was not possible to distinguish how the researcher’s intentions may have influenced the research outcomes. The remaining three studies were led by researchers who were not creators of the intervention and therefore could offer a reduced chance of this bias (Andras, 2012; Barakova et al., 2004; Owens et al., 2008).

Demographic

The studies were conducted within three countries: USA (LeGoff, 2004; LeGoff and Sherman, 2006), UK (Andras, 2012; Owens et al., 2008) and Netherlands (Barakova et al., 2004). The multinational nature of these studies allows support to the multicultural applicability of the intervention. However, when making this assumption it is important to note that these were western countries and the intervention may not be as well received in other cultures across the world.

All the participants were children with a clinical diagnosis of ASD. They were all within the age range of five to sixteen to represent the UK primary and secondary school ages - where this intervention could be implemented. Three of the studies used a primary aged sample (Andras, 2012; LeGoff and Sherman, 2006; Owens et al., 2008) another used an age range of eight to twelve (Barkova et al., 2004) and one study used an age range of six to sixteen (LeGoff, 2004). Andras (2012) was the only study to be implemented by a Teaching Assistant and had participants within a school setting. This study was subsequently awarded the highest score in WOE-C topic relevance. Andras (2012) explicitly researched how the LEGO® therapy intervention affects children in a school setting, therefore, the study matches ecologically and demographically well. The match of relevance and demographic strengthens the weight of this research and allows for a greater relevance of the findings compared to the other studies.

All the studies displayed a large gender disparity. The samples in four of the studies had over double the number of males to females (Andras, 2012; LeGoff, 2004; LeGoff and

Sherman, 2006; Owens et al., 2008) and with one study sample being all male (Barakova et al., 2004). This indicates a clear gender bias in the literature. However, it could be argued that the gender disparity could have been found due to ASD diagnosis being three times higher in males than females (Frederickson, 2008).

Sample Size

The sample size among the studies varied greatly in number, from 6 participants (Barakova et al., 2004) through to 117 (LeGoff and Sherman, 2006). To assess the power of these studies, reference was made to Cohen (1992) who suggests that group sizes for detection of effect size when comparing means at 0.5 significance level and when the power level is set at 0.8 are:

- Large – 23 participants
- Medium – 64 participants
- Small – 393 participants

According to these criteria only one of the group based studies was sufficiently powered (LeGoff and Sherman, 2006). The other two group based studies were underpowered (LeGoff, 2004; Owens et al., 2008). Therefore, caution should be taken when considering the effects that are stated by the underpowered studies. This is reflected in the WOE – A criteria (*Appendix 3*) where the two underpowered studies were scored lower.

Two of the studies utilised the single case design (Andras, 2012; Barakova et al., 2004). This design methodology allows the participant to act as their own control. This means

that within single case designs the sample size can be one. To allow this difference in methodological stance some separate criteria for WOE A and B were used to judge methodological quality. This was based on the work of Horner et al. (2005) and allowed both methodological stances to be held as equal.

Designs

There was a large variation in designs across the studies. Barakova et al. (2004) and Andras (2012) both explored how the LEGO® therapy intervention could encourage communication via small scaled within-subject design. The design does at first seem strong however it is hard to compare across groups due to the heterogeneity of the population. It should be noted that one study (Andras, 2012) was based in a school therefore the findings could be more relevant to the research question as it was ecologically relevant. A randomised control trial was used by Owens et al. (2008). They compared LEGO® therapy against Social Use of Language Programme (SULP, Rinaldi, 2004), with a control group, in improving social communication. Consequently, this study scored higher than the others in WOE-B (*Appendix 3*) due to its overall rigour and relevance of methodology.

A waiting list control design was used to explore the effects of LEGO® therapy and social communication by LeGoff (2004). The strength within this design lies in the fact that it considers confounding variables such as maturation. LeGoff and Sherman (2006) explored the long-term effects of LEGO® therapy by employing a thirty-six month pre

and post treatment series design. A comparison group was also used to strengthen the ability in distinguishing the effects of the intervention.

Procedure

Setting

The settings of the studies were varied. Four of the studies were conducted in a clinical setting (Barakova et al., 2004; LeGoff, 2004; LeGoff and Sherman, 2006; Owens et al., 2008). The intervention was facilitated by non-teaching staff members, usually the lead clinical researcher (LeGoff, 2004; LeGoff and Sherman, 2006; Owens et al., 2008). Barakova et al. (2004) used a robot for LEGO® therapy delivery, this consequently scored low on WOE-C (*Appendix 3*) due to the robot delivery method having a low relevance to the review topic. Andras (2012) was the only study to score high on WOE-C (*Appendix 3*) as the intervention was delivered in a school by a trained member of teaching staff. This holds great relevance to the research topic which is focusing on LEGO® therapy in a school based context.

Measures

Standardised Tests

The Gilliam Autism Rating Scale (GARS-SI, Gilliam 1995) was used by three of the studies as part of the pre and post measures (LeGoff and Sherman, 2006; LeGoff, 2004; Owens et al., 2008). The Vineland Adaptive Behavior Scales (VABS: Sparrow et al., 1984) was also used by Owens et al. (2008). Using these standardised tests allowed for measurements of outcomes to be justified confidently with statistical evidence. Out of all the studies that used the standardised assessments it was only

LeGoff (2004) who used just one source for his research. The other two studies (LeGoff and Sherman, 2006; Owens et al., 2008) obtained their standard measurements from more than one source. This is the reason why the study was given a lower rating in the WOE-A methodological quality section. This would further, imply that the multi sourced nature of the other two studies (LeGoff and Sherman, 2006; Owens et al., 2008) could be relied upon more confidently to measure the social communicational improvements that they suggest.

Observations

Three of the studies employed the use of observations as additional measures (LeGoff, 2004; Owens et al., 2008). Two of the studies used observation as their single method of measure (Andras, 2012; Barakova et al., 2004). Though, Andras (2012) did not use video observations unlike Barakova et al. (2004). This could reduce the reliability of the observation as observation bias could influence results. LeGoff (2004) was the only study to report inter-rater agreement. This allowed for a higher WOE-A for this study due to the strengthening of its methodological quality.

Findings

Positive effects of LEGO® therapy were found across all the studies with regards to social communication measures. The effect sizes of the studies can be seen in *Table 5*. The majority of the studies reported a large effect size. This is supporting evidence for the positive effect of LEGO® therapy on improving social communication. Effect sizes reported below were either given by the researcher (Barakova et al., 2014) who reported one effect size by combining all their participants from intervention compared

to baseline. LeGoff and Sherman, (2006); LeGoff, (2004); Owens et al. (2008) reported means and standard deviations therefore effect size could be calculated using the Pre-Post Standardised Mean Difference (PPSMD). This meant that these four studies could be compared to Cohen's d effect size indicators: 0.2 will be regarded as small, 0.5 medium and 0.8 large (Cohen, 1988). Andras (2012) did not provide any information of an effect size or enough information for one to be calculated.

14. Table 5

15. Outcomes and effect sizes of included studies

Study ID	Author	Outcome Measure	Participants Size	Effect Size (Cohen's d)	Overall Weight of Evidence
1	Barakova et al. (2014)	Observation of playing alone	n=6	Intervention compared to baseline (combined) = 1.0 (Large)	Low
2	Owens et al. (2008)	GARS-SI VABS -MAL	n=16 n=16	Intervention PPMSD = -0.2 (Small) Intervention PPMSD = -0.5 (Medium)	Medium
3	LeGoff and Sherman (2006)	VABS-SI GARS-SI	n=60 n=60	Intervention PPMSD = -1.5 (Large) Intervention PPMSD = -2.2 (Large)	Medium
4	LeGoff (2004)	Group 1: GARS SISC DSI Group 2: GARS SISC DSI	n=47 n=21	Intervention PPMSD = -0.9 (Large) Intervention PPMSD = 0.9 (Large) Intervention PPMSD = 1.2 (Large) Intervention PPMSD = -1.9 (Large) Intervention PPMSD = 1.18 (Large) Intervention PPMSD = 2.26 (Large)	Medium
5	Andras (2012)	Observation	n=8	Not enough information provided to calculate effect size	Medium

Effect sizes were interpreted according to Cohen's criteria, where '0.2' is considered a 'small' effect, '0.5' is regarded as a 'medium' effect and '0.8' is deemed be a 'large' effect (Cohen, 1988).

LeGoff and Sherman (2006) showed that LEGO® therapy resulted in improved social competence and fewer inappropriate behaviours. LeGoff (2004) also supported these findings. Their study showed an improvement in social competence after the LEGO® therapy intervention. They went on further to show that this effect was long term, of twenty-four weeks, in comparison to the control group which showed no improvement. Barakova et al. (2014) reported a large effect size in social interaction during the intervention. However, the study only reported one large effect size and did not include a considerable amount of analysis, which they said would be posted in a later paper. Owens et al. (2008) found that LEGO® therapy was a better intervention at increasing social interaction than the Sulp intervention. Andras (2012) showed that LEGO® therapy did have a positive impact on social interactions within her participants. However not enough data was reported to calculate an effect size. It would have been better if more rigorous analysis was conducted so that significance or effect sizes could be determined.

Section 4: Conclusions and Recommendations***16. Conclusion***

The five studies used in this review show a unanimous consensus on the positive effects of LEGO® therapy on improving social communication difficulties within children with ASD. There was promising evidence which suggested that LEGO® therapy could be a better way to improve social communication difficulties, than compared to other interventions (Owens et al., 2008). However, this was only one study with one other comparison intervention, further research would be needed to support such a claim.

It must be said that to generalise the effects of LEGO® therapy based on the evidence in this review would not be strictly correct. The majority of the evidence was rated as an overall weight of 'medium' therefore it does not provide the concrete evidence required to make such generalizations. The review has highlighted that appropriate sample size to power the studies was a key weakness in two of the group comparison studies (LeGoff, 2004; Owens et al., 2008) one group comparison study was sufficiently powered (LeGoff and Sherman, 2006). This would further suggest that the findings from those underpowered studies could not reflect the strength of effect that they were claiming. This would have consequences in attempting to generalise the effectiveness of the intervention. Finally, the gender disparity of the studies was significantly skewed towards males. To generalise these findings across genders could be inaccurate. The appeal of LEGO® could be far less significant within female participants therefore effecting the role LEGO® therapy would have on bringing about the positive social communication skills that it claims.

Recommendations

Lego® therapy does seem to be getting popular across schools as more schools are including the therapy as an intervention to support children with ASD (AET, 2012). The appeals of this naturalistic approach have ensured its popularity, however, the evidence base for this approach is still emerging. This review should allow Educational Psychologists to make a more informed choice about the appropriateness of LEGO® therapy to help their children with ASD. From this literature review it seems that the effects of LEGO® therapy can be substantial in helping children with a formal diagnosis of ASD with their communication and social skills. With this knowledge schools and Educational Psychologists can be confident that the intervention is likely to bring success to the child. Furthermore, the findings suggest that these improvements can be seen in both a school and home environment. This allows practitioners access to one intervention which can target the child's social skills in a variety of domains. This would be especially important to Educational Psychologists who work across a multitude of systems and settings.

The two areas that LEGO® therapy evidence seems to lack is the sample size and female participants. It would be hard to generalise the findings of this intervention without further research into these fields. LEGO® therapy is a relatively new and emerging intervention with promising initial evidence. It would be advisable for further research to address these highlighted gaps in the evidence to allow Educational Psychologists to make more informed choices.

References

- Alvarez, A. and Philips, A. (1998). The importance of play: A child psychotherapist's view. *Child Psychology and Psychiatry Review*, 3, 99–103.
- American Psychiatric Association (APA) (2013). *Diagnostic and Statistical Manual of Mental Disorders* (DSM-V-TR) (5th Edition) Washington, DC: American Psychiatric Association.
- Bodrova, E, and Leong, J. (2003). Building language and literacy through play. *Scholastic Early Childhood Today*, 18, (2), 34–8, 40–3.
- Baron-Cohen, S. (2006). The hyper-systemizing, assortative mating theory of autism. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 30(5), 865-872.
- Baron-Cohen, S. (2008). Autism, hypersystemizing, and truth. *The Quarterly Journal of Experimental Psychology*, 61(1), 64-75.
- Chambless, D.L, Baker, M.J, Baucom, D.H, Beutler, L.E, Calhoun, K.S, Crits-Cristoph P, Woody, S.R (1998). Update on empirically validated therapies, II. *The Clinical Psychologist*, 51, 3–16.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Department of Education (DoE, 2001). *Educational Provision and Support for Persons with Autistic Spectrum Disorders: the Report of the Task Force on Autism*. Bangor, Co Down: DoE.
- Department for Education and Skills (DfES, 2005). *Excellence and Enjoyment: Social and Emotional Aspects of Learning (guidance)*. Nottingham: DfES.
- Frederickson, N. (2002). Evidence-based practice and educational psychology. *Educational and Child Psychology*, 19, 3, 96-111.
- Frederickson, N. (2008). Educating children with autism. In N. Frederickson, A. Miller, and T. Cline (Eds.), *Educational Psychology* (pp. 142 – 156). London: Hodder Education.
- Goleman, D. (1996). *Emotional intelligence: Why it can matter more than IQ*. London: Bloomsbury.
- Goncu, A. (1999). Children's and researcher's engagement in the world. In A. Goncu. *Children's engagement in the world: Sociocultural perspectives*. New York: Cambridge.
- Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and

relevance of evidence. *Research Papers in Education*, 22(2), 213–228.

Guyatt, G., Sackett, D., Sinclair, J., Hayward, R., Cook, D., and Cook, R. (1995). Users' guide to the medical literature. IX. A method for grading health care recommendations. *Journal of the American Medical Association*, 274, 1800-1804.

Hart, J. E. (2011). Creating social opportunities for students with autism spectrum disorder in inclusive settings. *Intervention in School and Clinic*, 46, (5), 273-279.

Harrower, J. K., and Dunlap, G. (2001). Including children with autism in general education classrooms. *Behavior Modification*, 25, (5), 762 – 784.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., and Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165–179.

Howlin, P. R., Gordon, K., Pasco, G., Wade, A., and Charman, T. (2007). The effectiveness of Picture Exchange Communication System (PECS) training for teachers of children with autism: a pragmatic, group randomised controlled trial. *Journal of Child Psychology and Psychiatry*, 48, (5), 473–48.

Kazdin, A.E. (2010) Single-case research designs: Methods for clinical and applied settings. 2nd ed. *Oxford University Press*, New York

Kratochwill, T. R. (2003). Evidence-Based Practice: Promoting Evidence-Based Interventions in School Psychology. *School Psychology Quarterly*, 18(4), 389- 408.

LeGoff, D. (2004). Use of LEGO as a Therapeutic Medium for Improving Social Competence, *Journal of Autism and Developmental Disorders*, 34, (5), 557-571.

LeGoff, D., Krauss, G. and Allen Levin, S. (2010). LEGO®-based Play Therapy for Autistic Spectrum children. In Drewes, A. and Schaefer, C. *School-based play* (second edition) Hoboken: John Wiley and Sons.

Lord, C., and Pickles, A. (1996). Language level and nonverbal social-communicative behaviors in autistic and language delayed children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35, 1542–1550.

McConnell, S. R., Missall, K. N., Silberglitt, B., and McEvoy, M. A. (2002). Promoting social development in preschool classrooms. In M. E. Shinn, H. E. Walker, and G. Stoner. *Interventions for Academic and Behavior Problems II: Preventive and Remedial Approaches*. Bethesda, MD: National Association of School Psychologists.

National Research Council. (2001). *Educating children with autism*. Washington, DC: National Academy Press.

Owens, G. Granader, Y. Humphrey, A. and Baron-Cohen, S. (2008). LEGO Therapy and the Social Use of Language Programme: An Evaluation of Two Social Skills Interventions for Children with High Functioning Autism and Asperger Syndrome, *Journal of Autism Developmental Disorders*, 38, 1944-1957.

Spears, R., Tollefson, N., and Simpson, R. (2001). Usefulness of different types of assessment data in diagnosing and planning for a student with high-functioning autism. *Behavioral Disorders*, 26, (3), 227–242.

Stanton-Chapman, T. L., Denning, C. B., and Jamison, K R. (2012). Communication skill building in young children with and without disabilities in a preschool classroom. *The Journal of Special Education*, 46, (2), 78-93.

Timler, G. R., Olswang, L. B., and Coggins, T. E. (2005). Social communication interventions for preschoolers: Targeting peer interactions during peer group entry and cooperative play. *Seminars in Speech and Language*, 26, 170–180.

Walberg, J. L., and Craig-Unkefer, L. A. (2010). An examination of the effects of a social communication intervention on the play behaviors of children with autism spectrum disorder. *Education and Training in Autism and Developmental Disabilities*, 45, (1), 69-80.

17. Appendix 1 – List of all the excluded studies with exclusion reason

	EXCLUDED STUDIES	REASON FOR EXCLUSION
1	Kato, D., Hattori, K., Iwai, S., and Morita, M. (2012). Effects of collaborative expression using Lego® blocks, on social skills and trust. <i>Social Behavior and Personality: an international journal</i> , 40(7), 1195.	Criteria 4
2	Levy, F., Karel, K., and Soon, Y. (2016). Application of play therapy in registrar training. <i>Australian and New Zealand Journal of Psychiatry</i> , 50(4), 382–383.	Criteria 4
3	Evans, C. (2014) 'Lego Therapy' Club for children with Autism Spectrum Disorder. <i>Clinical Psychology Forum</i> .	Criteria 5

18. Appendix 2 – Mapping the field

LeGoff (2004)

Aims:	Participants	Country	Intervention	Design	Measures	Outcomes
<p><i>The use of LEGO® Therapy to try and improve the social competence of children with ASD.</i></p>	<p>Total of 47 participants all with a diagnosis of either ASD, Asperger’s or PDD-NOS. 34 males and 13 females. Age: 6-16 with a mean age of 10.</p>	<p>America</p>	<p>Delivery: Therapist in a clinic. 12 week course of 1 x per week of LEGO® Club (90 mins - group based) and 1 x 1 hour per week one-to-one.</p>	<p>Repeated measures, wait list control design.</p>	<p>School observations in unstructured situations with peers. Gilliam Autism Rating Scale social interaction subscale (GARS–SI).</p>	<p>No gender or age differences were found. Intervention group scored less on the social impairment scale compared to control. Significant improvement in motivation and ability in social contact with peers.</p>

LeGoff and Sherman (2006)

Aims:	Participants	Country	Intervention	Design	Measures	Outcomes
<p><i>A retrospective study of the long-term effect of LEGO® Therapy in improving the social competence of children with ASD.</i></p>	<p>Total of 117 participants. 60 participants all with a diagnosis of either ASD (n=26), Asperger's (n=27) or PDD-NOS (n=7). 49 males and 11 females. Mean age of 9:3. Control group with a similar diagnosis and demographic (n=57).</p>	<p>America</p>	<p>Delivery: Therapist in a clinic. 12 week course of 1 x per week of LEGO® Club (90mins) and 1 x 1 hour per week one-to-one. Run for 3 years.</p>	<p>Retrospective cohort study.</p>	<p>Compared Vineland Adaptive Behavior Scale socialization domain (VABS-SD) and Gilliam Autism Rating Scale social interaction subscale (GARS-SI) scores pre- and post-treatment with a matched comparison sample (N = 57).</p>	<p>LEGO® participants improved significantly more than the comparison subjects in social skills and reduction in ASD type behaviours (VABS-SD). LEGO® group also showed an improvement in their social competence in a natural setting.</p>

Owens, Granade, Humphrey and Baron-Cohen (2008)

Aims:	Participants	Country	Intervention	Design	Measures	Outcomes
<p><i>An evaluation of the effectiveness of LEGO® therapy and the Social Use of Use of Language Program (SULP) for children with high functioning Autism and Asperger's syndrome.</i></p>	<p>Total of 31 participants for both interventions. LEGO® Therapy: all with a diagnosis of either HFASD (n=5), Autistic Disorder (n=8) and ASD (n=3). Control group of no intervention n=16. Mean age of 8.2. All participants in mainstream education.</p>	<p>England</p>	<p>Delivery: therapist in a clinic. 18 week course of 1 x 1 hour per week of intervention in small groups.</p>	<p>Randomised control design</p>	<p>Observations. Gilliam Autism Rating Scale social interaction subscale (GARS-SI). Vineland Adaptive Behavior Scale socialization domain (VABS-SD) Child and parent satisfaction questioners.</p>	<p>Found that the most improvement on GARS-SI was in the LEGO® therapy group. Maladaptive behaviour decreased significantly more in the LEGO and SULP groups compared to the control group. Small but significant increase in duration of social interactions in LEGO® group, no change in SULP group. Children and parents more satisfied with LEGO® therapy than SULP.</p>

Andras (2012)

Aims:	Participants	Country	Intervention	Design	Measures	Outcomes	
<p><i>To explore the effects of LEGO® therapy on primary school aged children with ASD on their social interaction skills.</i></p>	<p>8 primary school children with a diagnosis of ASD.</p>	<p>England</p>	<p>Delivered by specialist teacher in mainstream primary school.</p>	<p>Within single subject design</p>	<p>Playground observations of social interactions.</p>	<p>Increase in social interactions between children after LEGO® therapy intervention.</p>	
	<p>Ages were 8-11 with a mean of 9.3.</p>		<p>10 week course of 1 x 45 minute session of LEGO® therapy per week.</p>				<p>Effect of increased interactions was maintained after therapy had ceased.</p>
	<p>7 male and 1 female.</p>		<p>Group-based intervention.</p>				

Barakova, Bajracharya, Willemsen, Lourens and Huskens (2014)

Aims:	Participants	Country	Intervention	Design	Measures	Outcomes
<p><i>An exploration into the effects of LEGO® therapy on children with ASD through a humanoid robot.</i></p>	<p>Total of 6 participants all with a diagnosis of ASD or PDD-NOS.</p>	<p>Netherlands</p>	<p>Delivered by a therapist in a clinic however session was mediated by a robot.</p>	<p>Multiple baseline design.</p>	<p>Observation protocol forms.</p>	<p>Significant increase in social initiations during the intervention, however was not observed after intervention.</p>
	<p>Participants aged between 8-12 with mean age of 10.3</p>		<p>4 week course of 1x 30 minute intervention per week.</p>			<p>Large effect of decreased occasions of children playing alone.</p>
	<p>All participants were male.</p>		<p>Group intervention.</p>			
	<p>2 participants had co-morbidities. 1 with a tic disorder and 1 with ADHD.</p>					<p>Increase in dyadic interactions between children and robot.</p>

19. Appendix 3: Weight of Evidence Criteria

WOE A – Quality of Methodology

As this study was looking at both comparison groups and single case designs it utilised two different coding protocols. To determine single case designs the Horner et al. (2005) coding protocol along with some scales from the Kratochwill, T.R. (2003) protocol form (to help with the Measures and Analysis section) was used. For the comparison groups design studies the Kratochwill, T.R. (2003) was the only protocol used.

Measures

<i>Weighting</i>	<i>Description</i>
<i>High (3 or more)</i>	<ul style="list-style-type: none"> • Study is able to report a reliability of .85 or above for primary outcomes which are relevant to this review question. • 2 or more methods of assessment. • 2 or more sources of assessment. • Uses either a well referenced, norm-referenced or standardised measure. • Reports the validity of all the measures used.
<i>Medium (2 or more)</i>	<ul style="list-style-type: none"> • Study is able to report a reliability of .85 for some of its primary outcomes. • Study used well referenced measures for some of its primary outcomes. • 2 or more methods of assessment. • 2 or more sources of assessment. • Criteria must be met for 75% of primary outcomes.
<i>Low (1 or more)</i>	<ul style="list-style-type: none"> • Study is able to report a reliability of .50 for some of its primary outcomes. • Study used well referenced measures for some of its primary outcomes. • Criteria must be met for 50% of primary outcomes.

Either A) Comparison group

<i>Weighting</i>	<i>Description</i>
<i>High (requires 3)</i>	<ul style="list-style-type: none"> • Study must have used an active comparison group. • Group equivalency must be established. • Evidence of counterbalancing of change agents as well as displaying equivalent mortality and low attrition at post.
<i>Medium (requires 3 or more)</i>	<ul style="list-style-type: none"> • A no intervention comparison group used (essential). • Counterbalancing of change agents. • Group equivalence established. • Equivalent mortality with low attrition.
<i>Low (2 or more)</i>	<ul style="list-style-type: none"> • Some sort of comparison group (essential) • Counterbalancing of change agents. • Group equivalence established. • Equivalent mortality with low attrition.

Or B) Single Case Designs

<i>Weighting</i>	<i>Description</i>
<i>High(all)</i>	<ul style="list-style-type: none"> • Multiple baseline design or 3 or more attempts to show intervention effect (Horner et al., 2005). • 3 or more data points for maintenance and generalisation phase.
<i>Medium (all)</i>	<ul style="list-style-type: none"> • 3 or more attempts to show intervention effect. • 3 or more data points for maintenance or generalisation phase.
<i>Low(all)</i>	<ul style="list-style-type: none"> • Intervention effect could have been shown less than 3 times. • Maintenance or generalisation data may not be included or less than 3 data points.

Analysis

<i>Weighting</i>	<i>Description</i>
<i>High (all)</i>	<ul style="list-style-type: none"> • Appropriate statistical analysis conducted. • All effect sizes can be calculated. • Sample size was sufficiently large. • Family/experimenter wise error was controlled for (where applicable).
<i>Medium (3 of 4)</i>	<ul style="list-style-type: none"> • Appropriate statistical analysis conducted. • Some effect sizes can be calculated. • Sample size was sufficiently large. • Pre and post measures were conducted.
<i>Low (2 of 3)</i>	<ul style="list-style-type: none"> • Appropriate statistical analysis conducted. • Some effect can be ascertained. • Sample size was sufficiently large.

WOE A was calculated by firstly assigning a score to each weight sub category:

High = 3

Medium = 2

Low = 1

The mean score was determined from the total score of the subcategories
(*Measures + Comparison Group/Single Case + Analysis* ÷ 3 = **WOE A**).

The final **WOE A** score was levelled using the following thresholds:

High = 2.6 to 3

**Medium = 1.5 to 2.5
below**

**Low = 1.4 or
below**

20. WOE B – Methodological Relevance

Either A) Comparison Group

<i>Weighting</i>	<i>Description</i>
<i>High (all)</i>	<ul style="list-style-type: none"> • Random assignment to either control or interventions group. • All groups must show pre and post measures.
<i>Medium (all)</i>	<ul style="list-style-type: none"> • Study uses comparison group. • All groups must show pre and post measures.
<i>Low (all)</i>	<ul style="list-style-type: none"> • All groups must show pre and post measures.

B) Single Case Design

<i>Weighting</i>	<i>Description</i>
<i>High (all)</i>	<ul style="list-style-type: none"> • Multiple baseline design or 3 or more attempts to show intervention effect (Horner et al., 2005).
<i>Medium (all)</i>	<ul style="list-style-type: none"> • 3 or more attempts to show intervention effect.
<i>Low (all)</i>	<ul style="list-style-type: none"> • Intervention effect could have been shown less than 3 times.

21. WOE B was calculated by assigning a score to each weight sub category:

High = 3

Medium = 2

Low = 1

22. WOE C – Topic relevance

<i>Weighting</i>	<i>Description</i>
<i>High (all)</i>	<ul style="list-style-type: none"> • Participants have a clinical diagnosis of ASD as recognised by the American Psychiatric Association. • Intervention is implemented by a school staff member with training on LEGO® therapy with roles of ‘Builder’ ‘Engineer’ and ‘Supplier’. • Intervention is delivered for 10 or more weeks with 1 session per week. • Pre and post measures of social communication are taken.
<i>Medium (all)</i>	<ul style="list-style-type: none"> • Participants are selected on researchers own diagnostic criteria for ASD. • Intervention is implemented by a therapist in a clinical setting that uses the roles of ‘Builder’ ‘Engineer’ and ‘Supplier’. • Intervention delivered for 5 or more weeks with 1 session per week. • Pre and post measures for social communication are described but not quantified.
<i>Low (all)</i>	<ul style="list-style-type: none"> • Participants have an assumed diagnosis of ASD e.g. suspicions from parents or teachers. • Intervention is implemented by someone other than a clinician or school staff member but still uses the roles of ‘Builder’ ‘Engineer’ and ‘Supplier’ e.g. parent or robot. • Participants social communication is not described.

WOE C was calculated by assigning a score to each weight sub category:

High = 3

Medium = 2

Low = 1

23. Overall Weight of Evidence (WOE D)

WOE D was calculated by firstly assigning a score to each weight sub category:

High = 3

Medium = 2

Low = 1

The mean score was determined from the total score of the subcategories ($WOE A + WOE B + WOE C \div 3 = WOE D$).

The final **WOE D** score was levelled using the following thresholds:

High = 2.6 to 3

**Medium = 1.5 to 2.5
below**

Low = 1.4 or below

Appendix 4 – Coding Protocols

Quality Indicators: Within Single-Subject Research (Horner et al., 2005)

Name of coder: X

Date: 21/1/17

Study number: 1

Full study:

Barakova, E. I., Bajracharya, P., Willemsen, M., Lourens, T., and Huskens, B. (2014). Long-term LEGO therapy with humanoid robot for children with ASD. *Expert Systems*, 32(6), 698–709.

Description of Participants and setting

	Yes	No
<i>Participants are described with sufficient detail to allow others to select individuals with similar characteristic; (e.g., age, gender, disability, diagnosis).</i>	✓	
<i>The process for selecting participants is described with replicable precision.</i>		✓
<i>Critical features of the physical setting are described with sufficient precision to allow replication.</i>	✓	

Dependent Variable

	Yes	No
<i>Dependent variables are described with operational precision.</i>		✓
<i>Each dependent variable is measured with a procedure that generates a quantifiable index.</i>		✓
<i>Measurement of the dependent variable is valid and described with replicable precision.</i>		✓
<i>Dependent variables are measured repeatedly over time.</i>	✓	
<i>Data are collected on the reliability or inter observer agreement associated with each dependent variable, and IOA levels meet minimal standards {e.g., IOA = 80%; Kappa = 60%}.</i>		✓

Independent Variable

	Yes	No
<i>Independent variable is described with replicable precision.</i>	✓	
<i>Independent variable is systematically manipulated and under the control of the experimenter.</i>		✓
<i>Overt measurement of the fidelity of implementation for the independent variable is highly desirable</i>		✓

Baseline

	Yes	No
<i>The majority of single-subject research studies will include a baseline phase that provides repeated measurement of a dependent variable and establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the independent variable did not occur.</i>	✓	
<i>Baseline conditions are described with replicable precision.</i>		✓

Experimental Control/internal Validity

	Yes	No
<i>The design provides at least three demonstrations of experimental effect at three different points in time.</i>		✓
<i>The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses).</i>		✓
<i>The results document a pattern that demonstrates experimental control.</i>	✓	

External Validity

	Yes	No
<i>Experimental effects are replicated across participants, settings, or materials to establish external validity.</i>		✓

Social Validity

	Yes	No
<i>The dependent variable is socially important.</i>	✓	
<i>The magnitude of change in the dependent variable resulting from the intervention is socially important.</i>	✓	
<i>Implementation of the independent variable is practical and cost effective.</i>		✓
<i>Social validity is enhanced by implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts.</i>		✓

[Adapted from Task Force on Evidence-Based Interventions in School Psychology, American Psychology Association, Kratochwill, T.R. (2003)]

A2 Multi-method (at least two assessment methods used)

Yes

No

N/A

Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)

Yes

No

N/A

Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

Yes validated with specific target group

In part, validated for general population only

No

Unknown/unable to code

C. Appropriate Statistical Analysis

Analysis

1 _____

Appropriate unit of analysis

Familywise/experimenter wise error rate controlled when applicable

[Adapted from Task Force on Evidence-Based Interventions in School Psychology, American Psychology Association, Kratochwill, T.R. (2003)]

Coding Protocol

Name of Coder: _____ x _____

Date: _____ 20/01/17 _____

Full Study Reference in proper format:

LeGoff, D. and Sherman, M. (2006) Long-term outcome of social skills intervention based on interactive LEGO play, *Autism*, 10, (4), 317-329.

Intervention Name (description of study): _____ LEGO® Therapy _____

Study ID Number: _____ 3 _____

Type of Publication:

Book/Monograph

Journal Article

Book Chapter

Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)

Completely randomized design

Randomized block design (between participants, e.g., matched classrooms)

Randomized block design (within participants)

Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)

Nonrandomized design

Nonrandomized block design (between participants)

Nonrandomized block design (within participants)

Nonrandomized hierarchical design

Optional coding for Quasi-experimental designs

A3. Overall confidence of judgment on how participants were assigned (select one of the following)

Very low (little basis)

Low (guess)

Moderate (weak inference)

High (strong inference)

Very high (explicitly stated)

N/A

Unknown/unable to code

B. Participants

Total size of sample (start of study): __ 117 __

Intervention group sample size: __ 60 __

Control group sample size: __ 57 __

C. Type of Program

- Universal prevention program
- Selective prevention program
- Targeted prevention program
- Intervention/Treatment**
- Unknown

D. Stage of Program

- Model/demonstration programs
- Early stage programs**
- Established/institutionalized programs
- Unknown

E. Concurrent or Historical Intervention Exposure

- Current exposure... Individual therapy Group therapy, Family therapy, Therapeutic aide, Speech-language therapy, Occupational therapy, Physical therapy, Psychiatric medications, Antipsychotics Psychostimulants SSRI**
- Prior exposure
- Unknown

2. Key Features for Coding Studies and Rating Level of Evidence/Support

(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A. Measurement (Estimating the quality of the measures used to establish effects)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes

- Yes
- No
- Unknown/unable to code**

A2 Multi-method (at least two assessment methods used)

- Yes VABS and GARS-SI**
- No
- N/A
- Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)

- Yes authors and independent child social workers.**
- No
- N/A
- Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

- Yes validated with specific target group
- In part, validated for general population only**
- No
- Unknown/unable to code

Overall Rating for measurement 1

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

B. Comparison Group

B1 Type of Comparison Group (Select one of the following)

- Typical intervention (typical intervention for that setting, without additions that make up the intervention being evaluated)
- Attention placebo
- Intervention element placebo
- Alternative intervention
- Pharmacotherapy
- No intervention
- Wait list/delayed intervention
- Minimal contact
- Unable to identify type of comparison

B2 Overall confidence of judgment on type of comparison group

- Very low (little basis)
- Low (guess)
- Moderate (weak inference)
- High (strong inference)
- Very high (explicitly stated)
- Unable to identify comparison group

B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc were counter-balanced across intervention)

- By change agent
- Statistical (analyse includes a test for intervention)
- Other
- Not reported/None

B4 Group equivalence established (select one of the following)

- Random assignment
- Posthoc matched set
- Statistical matching
- Post hoc test for group equivalence

B5 Equivalent mortality

- Low attrition (less than 20 % for post)
- Low attrition (less than 30% for follow-up)
- Intent to intervene analysis carried out?
Findings_____

Overall rating for Comparison group 2

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence

C. Appropriate Statistical Analysis

Analysis

1_____

- Appropriate unit of analysis

- Familywise/experimenter wise error rate controlled when applicable
- Sufficiently large N

Overall rating for Analysis 2

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence