How Effective is Shape Coding at Improving Understanding and use of Grammar in School-Aged Children with Developmental Language Disorder?

Sophie Brown and Sarah Wright

Abstract

Language plays a crucial role in children's development and daily lives. However, it is estimated that 7.58% of children in the United Kingdom are entering mainstream schools with a language disorder, without a known cause, that hinders their academic progress. This is often called developmental language disorder (DLD), although there is contention within the literature and amongst practitioners about the disorder's diagnostic criteria, name and aetiology. Nevertheless, DLD has a significant negative impact on children's daily functioning and life outcomes. One aspect of language that children with DLD often struggle with is grammar. Susan Ebbels, a speech and language therapist, created the shape coding system to visually and explicitly teach children with DLD grammar. Shape coding is built upon other similar systems to enable more complex grammar to be represented and is theoretically underpinned by the procedural deficit hypothesis (discussed in this paper). A systematic literature search was conducted with a targeted search question to explore "how effective is shape coding at improving understanding and use of grammar in school-aged children with developmental language disorder?". After screening and exclusions, five studies were identified and evaluated. Due to methodological limitations and variation, it was impossible to conclude whether shape coding is an effective intervention and if it would be effective for any area of grammar. Consequently, further large-scale research is necessary to determine shape coding's effectiveness. Implications for schools are considered, such as whether shape coding is only appropriate and cost-effective in specialist settings delivered by speech and language therapists.

How effective is shape coding at improving understanding and use of grammar in school-aged children with developmental language disorder?

The ability to understand and use language plays a crucial role in children's development and is considered at "*the heart of children's learning*" (Ebbels et al., 2024, p. 804). It is estimated that 7.58% of children entering mainstream schools in Year 1 in the United Kingdom (UK; approximately two children in every class of 30 pupils) experience a language disorder (LD), without a known cause, that hinders their academic progress (Norbury et al., 2016). There is inconsistency in the terminology and criteria used to describe children with LDs (Bishop, 2017; Bishop et al., 2016, 2017). In the last years the literature has shifted from using the term 'specific language impairment' (SLI) to 'developmental language disorder' (DLD) to describe children with language difficulties that negatively impact daily life and are not associated with another condition, (e.g., Autism or Down Syndrome) while LD is considered an umbrella term (Bishop et al., 2017). DLD significantly impacts individuals' daily functioning and life outcomes, such as academic achievement, employment, and mental health (Clegg et al., 2005; Conti-Ramsden et al., 2018; Law et al., 2009; Yew & O'Kearney, 2013). This underscores the importance of early intervention in helping school-aged children with DLD enhance their language skills and improve their long-term outcomes. One potential intervention is shape coding (SC).

Shape Coding

The national curriculum sets guidance regarding when grammatical concepts should be first introduced but not when they should be completely understood and how they should be taught or adapted for pupils with DLD (Department for Education, 2014a, 2014b). In this context, SC was developed by Susan Ebbels, a speech and language therapist (SLT) at Moor House School & College (a non-maintained special school in the UK) to provide a visual way to teach grammar to children with DLD (Moor House School & College, 2023). Ebbels (2007) built upon features of the colour pattern

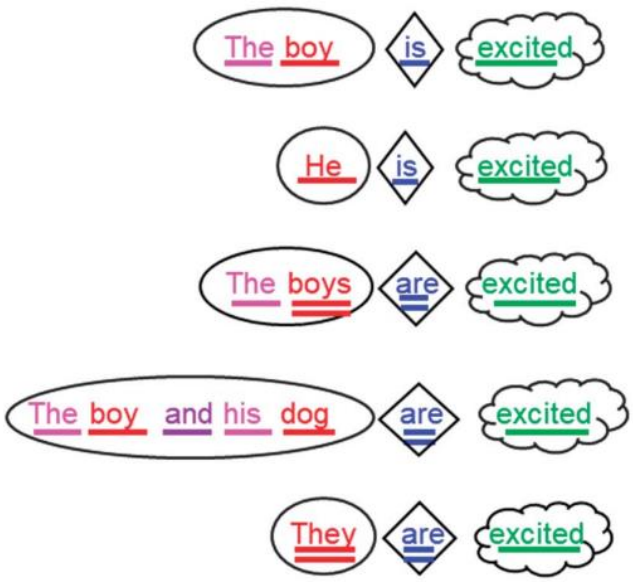
scheme (Lea, 1970) and colourful semantics (Bryan, 1997) to represent more complex grammar. SC is comprised of four main components to represent grammar:

- Shapes demonstrate the syntactic structure of a sentence by grouping words or phrases that answer different questions, such as "who" and "what doing". For example, ovals answer "who" or "what" questions, and hexagons answer "what doing".
- Colours demonstrate word classes (e.g., nouns are red, and verbs are blue), and in some shapes, a particular colour word is required (e.g., a red word is used in ovals).
- Lines highlight singular (underlined by a single line) and plural (underlined by a double line) nouns and verbs, and thus noun-verb agreement.
- Arrows demonstrate verb tenses and aspect (continuous or complete) with their direction. For example, present tense verbs are underlined in blue with a vertical arrow pointing down in the middle of the line, whereas for past tense, the arrow is pointing down at the left-hand end of the line.

See Figures 1 and 2 for examples of the SC system in use.

Figure 1

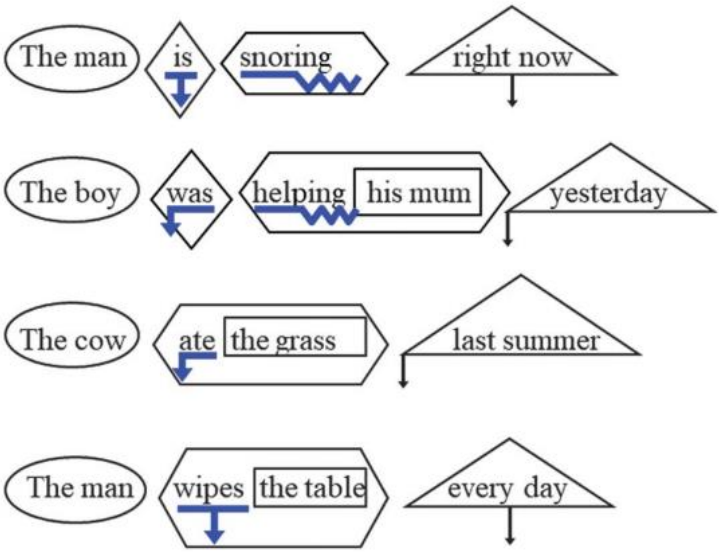
Demonstrating noun-verb agreement using the shape coding system



Note. This figure is from Balthazar et al. (2020).

Figure 2

Coding of verb tenses and time using the shape coding system.



Note. This figure is from Balthazar et al. (2020).

The first step in using SC is deciding with which areas of grammar a child needs support. Next, the order of teaching grammatical rules should be based on what would most benefit the child functionally. SC is first introduced using grammatical structures the child can produce and understand. Once the child understands the foundational elements, they can be taught new grammatical targets. A SC session can last approximately 15 to 30 minutes (Ebbels et al., 2007, 2014, 2024; Moor House School & College, 2023) and involves multiple steps:

1. The adult explicitly explains the target rule using SC and may use small figures or objects on the shapes to help act out and explain the meaning of the sentence.
2. The adult models the target rule using SC again, and the child is encouraged to act out or explain the sentence to demonstrate comprehension and allow the opportunity for adult feedback if required.

3. The adult encourages the child to produce the target rule orally or written using SC. If the child is accurate five consecutive times, the templates are removed so the child can try to produce the target without visual support. If the child makes two errors, the templates are reintroduced with acting out.
4. The adult and child take turns producing sentences with the target rule until the desired number of correct productions are made.
5. Once the child has internalised the rule (able to produce and comprehend it), the focus is on generalising to situations where the target rule is likely relevant. This also involves the adult making errors for the child to notice and correct.

Throughout sessions, errors are corrected and discussed. To use SC, The Moor House Research & Training Institute (2024) offer courses for SLTs and education professionals. The 'part one' course is self-paced online (approximately seven to 14 hours), costing £149 per person. Two more courses run twice a year via Zoom across two days to build confidence in using SC (practical applications course, £150 per person) and extend knowledge of SC for more complex language (part two course, £200 per person). Some free resources are on the SC website (e.g., a spreadsheet to identify grammatical targets and downloadable shapes; Moor House School & College, 2023). Additionally, there is a web app (£1.99 a month or £14.99 annually) and an app for tablets or phones to support the system if desired (£19.99 per device or a free 'lite' app version).

Theoretical Underpinnings

Ebbels (2007) indicated that two hypotheses underpin SC. The first is that children with DLD have difficulty learning language implicitly and, therefore, require explicit teaching of grammatical rules. The second is that children with DLD have visual strengths; thus, using visual coding optimises that strength to learn language explicitly. While Ebbels (2007) drew upon research to support the first hypothesis (e.g., Bishop et al., 2006), she did not provide evidence that children with DLD have visual strengths. Instead, Ebbels (2007) referenced that several other interventions use visuals and

colour coding to teach language (e.g., colourful semantics; Bryan, 1997), which themselves do not all have significant evidence bases (Law et al., 2012b). Therefore, SC was created and implemented while its theoretical underpinnings were still in its infancy.

Nevertheless, Ebbels and her colleagues (Calder et al., 2021; Ebbels et al., 2014) began endorsing Ullman and Pierpont's (2005) procedural deficit hypothesis (PDH) to explain why SC would be effective. The PDH proposes that many individuals with DLD have "abnormalities" in the brain structures involved with procedural memory, which results in difficulties with learning grammar implicitly (Ullman & Pierpont, 2005, p.399). The procedural memory system supports learning new skills, habits and rules (e.g., grammatical rules) implicitly through repeated stimulus exposure and practice. Whereas the declarative memory system depends on different brain structures and is responsible for learning, representing and using knowledge about facts and events and explicit retrieval and recognition of information (Calder et al., 2021; Ullman & Pierpont, 2005). Ullman and Pierpont (2005) suggest that due to a deficit in procedural memory and neuroplasticity, individuals with DLD experience declarative memory compensation. For example, they can learn grammatical rules explicitly as facts, such as recalling the rule "-ed" is added to the end of a verb when an event has already happened, rather than implicitly knowing the rule (Ullman & Pierpont, 2005). Therefore, SC aligns with the recommendations from the PDH, which is to explicitly teach children grammatical rules using the declarative system.

Several empirical studies support the PDH (Lum et al., 2012, 2014, 2015) and highlight the visual strengths of children with DLD, supporting why SC would be effective (Calder et al., 2021; Ebbels et al., 2014). For example, Lum et al. (2012) measured working, declarative and procedural memory in 51 children with DLD and 51 'typically developing' (TD) children with a mean age of 10. Compared to TD children, the children with DLD had impaired procedural memory. In contrast, their declarative memory for visual and verbal information was intact (once working memory and language skills were controlled for). Children with DLD's visuo-spatial short-term memory was also

intact. Additionally, grammatical abilities were associated with procedural memory in TD children but with declarative memory in children with DLD. Therefore, the findings support the predictions of the PDH and highlight the relative visual strengths of children with DLD. Therefore, SC has theoretical underpinnings, which are informed by research.

Additionally, whilst other theories explain DLD, resulting in different implications for practice that may not endorse SC, the PDH accounts best for the heterogeneity of DLD (Ebbels et al., 2014; Ullman & Pierpont, 2005). There is debate within the literature regarding the aetiology of and theory underlying DLD, often categorised into two theoretical perspectives (Ebbels et al., 2014; Ullman & Pierpont, 2005). For instance, linguistic theories posit that people with DLD have language-specific deficits, creating difficulties with grammar. Contrarily, processing theories propose that DLD is caused by processing difficulties that impact language. However, Ullman and Pierpont (2005) highlight that while other theories can explain certain aspects of DLD, they cannot account for the heterogeneity of the group and range of difficulties. Meanwhile, PDH explains that some of the group's heterogeneity is due to variations in the affected brain structures within the procedural memory system and to what degree they are affected. Therefore, the PDH, which supports SC, is the most coherent theory to explain DLD and thus may have the most relevant recommendations for practice (i.e., explicit grammar teaching, supported visually).

Impact and effectiveness of SC

A systematic literature search was conducted to ascertain whether the evidence suggests that SC is effective at improving understanding and use of grammar in school-aged children with DLD (see Appendix A for the search strategy and study flow and Appendix B for the quality appraisal table). Five studies were identified, all conducted in the United Kingdom and spanning over 23 years. Sample sizes ranged between three to 27, including participants aged eight to 16 years old who attended specialist schools in the UK due to language impairments (some had specific diagnoses of DLD). The study interventions were all carried out by SLTs. Hong et al.'s (2019) mixed methods

appraisal tool indicated that the studies were overall of good quality as they met most criteria. It is important to note that Ebbels, the creator of SC, was the first author of all five studies and gains financially from the SC courses and resources sold. Ebbels explicitly declares this conflict of interest in some studies (Ebbels et al., 2014, 2024). She is also on the editorial board of two journals: the 'International Journal of Language and Communication Disorders', which published Ebbels et al. (2014), and 'Child Language Teaching and Therapy', which published Ebbels (2007; Moor House School & College, 2023). It is also important to note that such potential conflicts of interest are not unique to SC.

Two studies used randomised control trials (RCTs) to investigate the effectiveness of SC (with children aged 11 to 16), which is often considered the 'gold standard' design to establish whether an intervention is effective (Snowling & Hulme, 2011). For example, Ebbels et al. (2014) randomly assigned participants (n = 27) to either the therapy group or 'waiting controls' to establish whether SC improved comprehension of coordinating conjunctions. Since the children were randomly assigned, pre-existing biases are reduced, and any measurable differences are attributed to the intervention effects. In phase one, the therapy group received SC across eight weeks (thirty minutes per week), whereas in phase two, they received no further intervention, and the waiting controls received SC. All participants were assessed on their comprehension of coordinating conjunctions pre- and post-phase one, post-phase two and at a four-month follow-up. They found that the therapy group made statistically significant gains compared to the waiting controls in phase one and maintained this gain during phase two. Additionally, the waiting controls made similar gains to that of the original therapy group once they received SC. This suggests that SC is an effective intervention. Ebbels et al. (2007) found similar results in their small-scale RCT (n = 14), where children who received SC over nine weeks made statistically significant gains in using verb argument structures, which was maintained at a three-month follow-up. Therefore, two studies using the 'gold standard' RCT design suggest that SC effectively improves children's understanding and use of grammar.

The remaining three studies used multiple baseline designs, which is appropriate for the small sample sizes (Levin & Ferron, 2021). Ebbels and van der Lely (2001) found in their pilot study that all participants (*n* = 4, aged 11 to 12) made progress with passives and with comprehension and production of 'wh' questions due to SC, albeit the significance and maintenance of the progress varied between individuals. The multiple baseline design enabled inferring the causal relationship between SC and improvement in performance, as a baseline in measures for both targets was established before SC. The baseline testing was repeated once per week for four weeks, highlighting that any change was due to SC (rather than other external variables). Ebbels (2007) built upon this research with some of the same participants (*n* = 3) in two studies and a third study with a class (*n* = 9, aged 11 to 13). Ebbels (2007) reported similar results to those of Ebbels and van der Lely (2001), where significant progress was found for some children but not all. More recently, Ebbels et al. (2024) employed a multiple baseline design with eight children (aged eight to 10) and found that scores on probe tests (tests to assess the children's understanding of the grammatical target) post-SC were significantly higher than during the baseline phase. They also used the design to explore how many teaching episodes per target per session is optimal (i.e., 10, 20 or 30). It was concluded that SC is effective and efficient, particularly with high teaching episode rates.

While the five studies employed well-regarded designs to explore SC's effectiveness, several limitations hinder confidently concluding whether SC is effective. Firstly, the most considerable limitation recognised in each study is the small sample size. Ebbels et al. (2014) highlighted that this limited power in their analyses, so they could not explore whether some findings were statistically significant. Moreover, they cautioned the importance of not under- or over-interpreting the data due to the small sample, as a large sample could influence the findings. Additionally, Ebbels et al. (2014, 2024) noted the small sample size limits the generalisation of the results to a broader population of children with DLD, as the participants attended the same educational setting and interacted with specific SLTs. Furthermore, Ebbels et al. (2014) stated some participants had previous knowledge of SC, which could affect their response to intervention. The studies all recommended a need for larger,

more diverse samples. However, interestingly, they did not document within their studies the socio-economic status or the ethnicity of their participants (except for Ebbels et al., 2007) to allow a better understanding of their sample. Moreover, there was no specific comment within their limitations on the imbalance of gender in all five of the studies, with boys dominating (who may respond to SC differently and present with DLD differently). Thus, further research with larger, more diverse sample sizes is necessary to determine whether SC is an effective intervention.

Secondly, it is challenging to compare or merge the findings of the studies because of key methodological differences (e.g., area of grammar and intervention period), and therefore, there is a lack of a coherent evidence base. For example, each study explored different areas of grammar and, therefore, cannot be combined as evidence to support SC being effective at improving grammar generally, but just for that specific target. Moreover, given there is only one study on each different area of grammar, there needs to be more studies per area to support that SC is effective at improving that target. Another difference between the studies is the intervention period and, thus, the total number of sessions needed for an impact. For example, Ebbels et al.'s (2024) intervention ran over 20 weeks, whereas Ebbels et al.'s (2014) intervention was only over eight weeks. Moreover, there was variance in the number of SLTs involved in each study and the detail of their experience, which may have impacted the effectiveness of the intervention. For example, Ebbels et al. (2024) reported that the SLT had three years of experience using SC. Ultimately, there are substantial methodological differences between the studies, mostly notably in the area of grammar explored, preventing generalisation of the positive findings and establishing a coherent evidence base to assert that SC is effective.

Summary and Implications

SC has a theoretical underpinning, Ullman and Pierpont's (2005) PDH, suggesting it might be effective. However, due to methodological limitations (small sample size) and variation (e.g., area of grammar targeted) in only five studies, it is impossible to conclude whether SC is an effective

intervention and if it would be effective for any area of grammar. Moreover, SC's optimal frequency (e.g., how many times a week over a period) and dosage (e.g., session length and number of teaching episodes per target per session) are not established. Nevertheless, Ebbels (2014) and Law et al. (2012a) highlight an important notion: while the evidence base is not strong, it does not mean SC is ineffective. Instead, there is just not enough evidence yet to conclude how effective it is and in what conditions. Consequently, further research is necessary to determine SC's effectiveness so that professionals and children spend their time on effective and efficient interventions (Ebbels, 2014; Ebbels et al., 2024).

Given the limited evidence, schools should consider how relevant SC is to their setting and its cost-effectiveness. For example, in all the studies, SC was delivered in specialist settings to children who attended the setting due to language impairment (whether diagnosed with DLD or not). Therefore, whether SC is effective for children with further additional needs or those in mainstream settings with language difficulties is unknown. Moreover, SC was delivered by SLTs in all the studies. Moor House Research & Training Institute (2024) offer courses to SLTs and professionals working within education. However, it is unclear if that includes teachers and teaching assistants and how that may impact the delivery and effectiveness of SC, as SLTs are likely more well-versed in delivering language interventions. Given that all three courses are just under £500 per person, SC is a significant investment before the additional annual costs of the apps. Therefore, SC may only be appropriate for specialist settings, yet still a considerable investment when there is currently a limited evidence base.

References

Balthazar, C. H., Ebbels, S., & Zwitserlood, R. (2020). Explicit grammatical intervention for developmental language disorder: Three approaches. *Language, Speech, and Hearing Services in Schools*, *51*(2), 226–246. https://doi.org/10.1044/2019_LSHSS-19-00046

Bishop, D. V. M. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language & Communication Disorders*, *52*(6), 671–680. https://doi.org/10.1111/1460-6984.12335

Bishop, D. V. M., Adams, C. V., & Rosen, S. (2006). Resistance of grammatical impairment to computerized comprehension training in children with specific and non-specific language impairments. *International Journal of Language & Communication Disorders*, *41*(1), 19–40. https://doi.org/10.1080/13682820500144000

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & and the CATALISE-2 consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Consortium, C. (2016). CATALISE: A multinational and multidisciplinary delphi consensus study. Identifying Language Impairments in Children. *PLOS ONE*, *11*(7), e0158753. https://doi.org/10.1371/journal.pone.0158753

Bryan, A. (1997). Colourful semantics: Thematic role theory. In S. Chiat, J. Law and J. Marshall (Eds.), Language disorders in children and adults: Psycholinguistic approaches to therapy (pp. 143-161). London: Whurr Publishers Ltd.

Calder, S. D., Claessen, M., Ebbels, S., & Leitão, S. (2021). The efficacy of an explicit intervention approach to improve past tense marking for early school-age children with developmental language disorder. *Journal of Speech, Language, and Hearing Research*, *64*(1), 91–104. https://doi.org/10.1044/2020_JSLHR-20-00132

Calder, S. D., Claessen, M., & Leitão, S. (2018). Combining implicit and explicit intervention approaches to target grammar in young children with developmental language disorder. *Child Language Teaching and Therapy*, *34*(2), 171–189. https://doi.org/10.1177/0265659017735392

Clegg, J., Hollis, C., Mawhood, L., & Rutter, M. (2005). Developmental language disorders – a follow-up in later adult life. Cognitive, language and psychosocial outcomes. *Journal of Child Psychology and Psychiatry*, *46*(2), 128–149. https://doi.org/10.1111/j.1469-7610.2004.00342.x

Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language & Communication Disorders*, *53*(2), 237–255. https://doi.org/10.1111/1460-6984.12338

Department for Education. (2014a, July). *National curriculum in England: English programmes of study.* Retrieved April 1, 2025, from https://www.gov.uk/government/publications/national-curriculum-in-england-english-programmes-of-study/national-curriculum-in-england-english-programmes-of-study#:~:text=Spoken%20languagePupils

Department for Education. (2014b, July). *English appendix 2: Vocabulary, grammar and punctuation*. Retrieved April 1, 2025, from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/335190/English_Appendix_2_-_Vocabulary_grammar_and_punctuation.pdf

Ebbels, S. (2007). Teaching grammar to school-aged children with specific language impairment using Shape Coding. *Child Language Teaching and Therapy*, *23*(1), 67–93. https://doi.org/10.1191/0265659007072143

Ebbels, S. (2014). Effectiveness of intervention for grammar in school-aged children with primary language impairments: A review of the evidence. *Child Language Teaching and Therapy*, *30*(1), 7–40. https://doi.org/10.1177/0265659013512321

Ebbels, S. H., Gadd, M., Nicoll, H., Hughes, L., Dawson, N., Burke, C., Calder, S. D., & Frizelle, P. (2024). The effectiveness of individualized morphosyntactic target identification and explicit intervention using the shape coding system for children With developmental language disorder and the impact of within-session dosage. *Language, Speech & Hearing Services in Schools*, *55*(3), 803–837. https://doi.org/10.1044/2024_LSHSS-23-00098

Ebbels, S. H., Marić, N., Murphy, A., & Turner, G. (2014). Improving comprehension in adolescents with severe receptive language impairments: A randomized control trial of intervention for coordinating conjunctions. *International Journal of Language & Communication Disorders*, *49*(1), 30–48. https://doi.org/10.1111/1460-6984.12047

Ebbels, S. H., van, der L. H. K. J., & Dockrell, J. E. (2007). Intervention for verb argument structure in children with persistent SLI: A randomized control trial. *Journal of Speech, Language, and Hearing Research*, *50*(5), 1330–1349. https://doi.org/10.1044/1092-4388(2007/093)

Ebbels, S., & van der Lely, H. (2001). Meta-syntactic therapy using visual coding for children with severe persistent SLI. *International Journal of Language & Communication Disorders*, *36*(S1), 345–350. https://doi.org/10.3109/13682820109177909

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions* (2nd ed.). John Wiley & Sons.

Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2019). Improving the content validity of the mixed methods appraisal tool: A modified e-delphi study. *Journal of Clinical Epidemiology*, *111*, 49-59. https://doi.org/10.1016/j.jclinepi.2019.03.008

Lea, J. (1970). *The colour pattern scheme: a method of remedial language teaching.* Moor House School.

Law, J., Lee, W., Roulstone, S., Wren, Y., Zeng, B., & Lindsay, G. (2012a). *"What works": Interventions for children with speech, language and communication needs*. Retrieved October 31, 2024, from

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/556912/DFE-RR247-BCRP10.pdf

Law, J., Lee, W., Roulstone, S., Wren, Y., Zeng, B., & Lindsay, G. (2012b). *"What works": Interventions for children with speech, language and communication needs: Technical annex.* Retrieved October 31, 2024, from https://assets.publishing.service.gov.uk/media/5a809806ed915d74e33fb502/DFE-RR247-BCRP10a.pdf

Law, J., Rush, R., Schoon, I., & Parsons, S. (2009). Modeling developmental language difficulties from school entry into adulthood: Literacy, mental health, and employment outcomes. *Journal of Speech, Language, and Hearing Research*, *52*(6), 1401–1416. https://doi.org/10.1044/1092-4388(2009/08-0142)

Levin, J. R., & Ferron, J. M. (2021). Different randomized multiple-baseline models for different situations: A practical guide for single-case intervention researchers. *Journal of School Psychology*, *86*, 169–177. https://doi.org/10.1016/j.jsp.2021.03.003

Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex*, *51*, 1–10. https://doi.org/10.1016/j.cortex.2013.10.011

Lum, J. A. G., Conti-Ramsden, G., Page, D., & Ullman, M. T. (2012). Working, declarative and procedural memory in specific language impairment. *Cortex*, *48*(9), 1138–1154. https://doi.org/10.1016/j.cortex.2011.06.001

Lum, J. A. G., Ullman, M. T., & Conti-Ramsden, G. (2015). Verbal declarative memory impairments in specific language impairment are related to working memory deficits. *Brain and Language*, *142*, 76–85. https://doi.org/10.1016/j.bandl.2015.01.008

Moor House Research & Training Institute. (2024). *Shape coding courses.* Retrieved October 31, 2024, from https://www.moorhouseinstitute.com/shape-coding-courses/

Moor House School & College. (2023). *Shape coding by Susan Ebbels.* Retrieved October 31, 2024, from https://shapecoding.com/

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, *57*(11), 1247–1257. https://doi.org/10.1111/jcpp.12573

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Snowling, M. J., & Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology*, *81*(1), 1–23. https://doi.org/10.1111/j.2044-8279.2010.02014.x

Ullman, M. T., & Pierpont, E. I. (2005). Specific Language Impairment is not Specific to Language: The Procedural Deficit Hypothesis. *Cortex*, *41*(3), 399–433. https://doi.org/10.1016/S0010-9452(08)70276-4

Yew, S. G. K., & O'Kearney, R. (2013). Emotional and behavioural outcomes later in childhood and adolescence for children with specific language impairments: Meta-analyses of controlled prospective studies. *Journal of Child Psychology and Psychiatry*, *54*(5), 516–524. https://doi.org/10.1111/jcpp.12009

Appendix A: Systematic Search Strategy and Study Flow

An initial scoping search was conducted using Google Scholar and PsychInfo with the terms 'shape coding' and 'language'. The PICOS framework (Higgins et al., 2019) was then applied to identify key search terms (see Table 1) and formulate the research question and inclusion and exclusion criteria (see Table 2). The research question was "how effective is shape coding at improving understanding and use of grammar in school-aged children with developmental language disorder?". The search terms were used in the following online databases: PyschInfo, Education Resources information Centre (ERIC) and Scopus. The grey literature was searched using the ProQuest Dissertation and Theses Global database.

Table 1

Search Terms Applied to the PICOS Framework

Population	child* OR school-age OR pupil* OR student* OR adolescent* OR teen* OR "young people" OR youth
Intervention	AND "shape coding"
Comparison	
Outcome	AND language
Setting	OR school*

Table 2

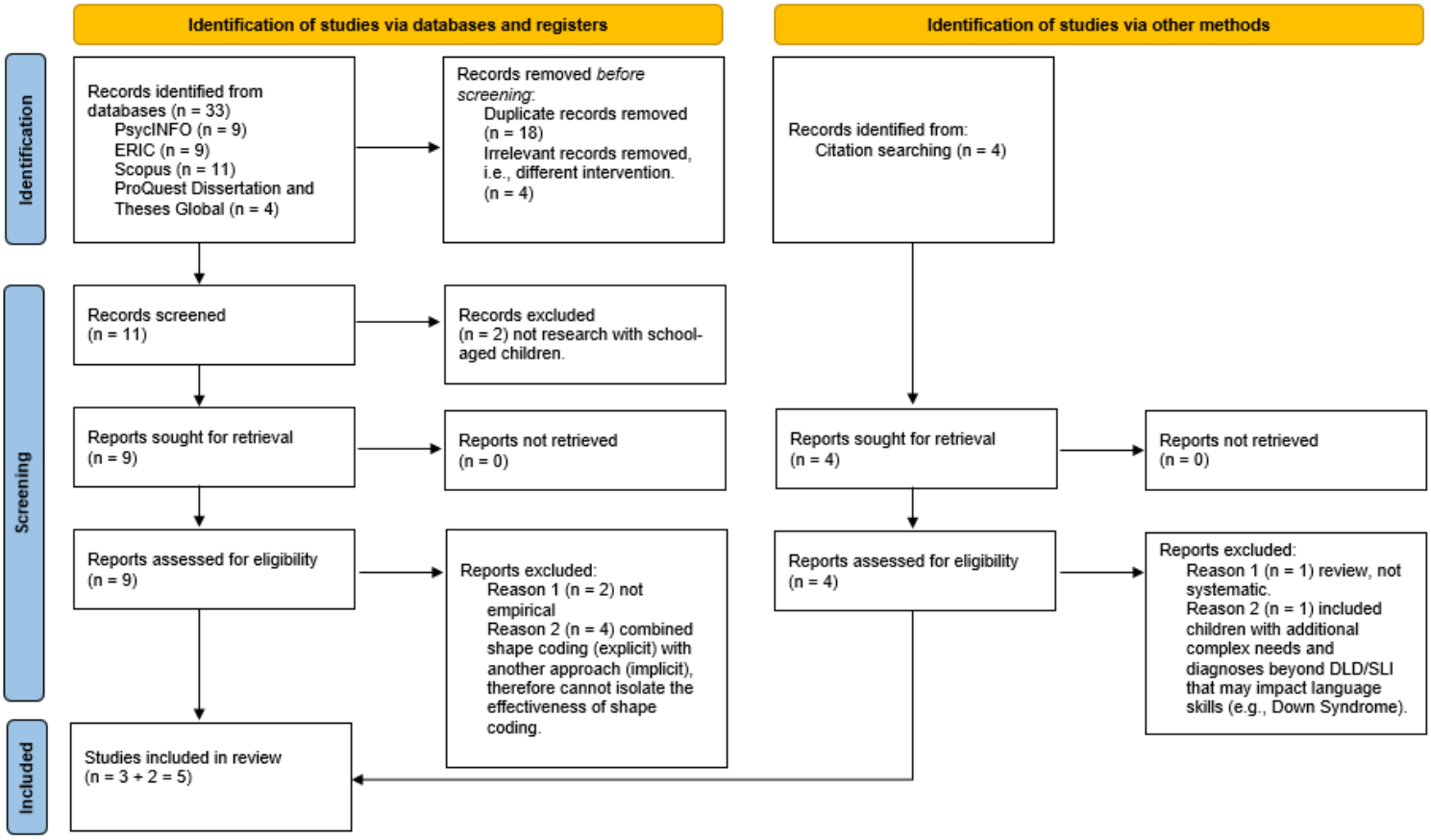
Inclusion and Exclusion Criteria for Studies

Inclusion criteria	Specified as shape coding. Participants aged 16 or under. Participants identified as having a language impairment, SLI or DLD. School-based intervention. Full text available. Empirical study.

	Presence of a comparison or control group (this could include treatment as usual and/or wait list. Assessed understanding or use of grammar. Peer reviewed article, or thesis or dissertation. Published in English.
Exclusion criteria	Not specified as shape coding. Participants aged 17 years or older. Participants identified as having a language impairment due to a neurodevelopmental condition such as Autism or Down Syndrome. Review papers. Non-empirical papers (e.g., reviews that are not systematic, opinion pieces, theoretical papers, book chapters, qualitative) Did not include a control comparison group. Did not include grammar related outcomes. Not conducted in a school setting. Full text not available Published in a language other than English

Figure 1 demonstrates the PRISMA flow diagram from Page et al. (2021) of how studies were identified, screened and excluded or included, resulting in the final five studies.

Figure 1


Identification of studies via databases and registers
Identification of studies via other methods
Identification
Screening
Included
Records identified from databases (n = 33)
PsycINFO (n = 9)
ERIC (n = 9)
Scopus (n = 11)
ProQuest Dissertation and Theses Global (n = 4)
Records removed *before screening*:
Duplicate records removed (n = 18)
Irrelevant records removed, i.e., different intervention. (n = 4)
Records identified from:
Citation searching (n = 4)
Records screened (n = 11)
Records excluded (n = 2) not research with school-aged children.
Reports sought for retrieval (n = 9)
Reports not retrieved (n = 0)
Reports sought for retrieval (n = 4)
Reports not retrieved (n = 0)
Reports assessed for eligibility (n = 9)
Reports excluded:
Reason 1 (n = 2) not empirical
Reason 2 (n = 4) combined shape coding (explicit) with another approach (implicit), therefore cannot isolate the effectiveness of shape coding.
Reports assessed for eligibility (n = 4)
Reports excluded:
Reason 1 (n = 1) review, not systematic.
Reason 2 (n = 1) included children with additional complex needs and diagnoses beyond DLD/SLI that may impact language skills (e.g., Down Syndrome).
Studies included in review (n = 3 + 2 = 5)

Appendix B: Quality Assurance of Studies Included

The final five studies included in the systematic review were appraised for their quality using the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2019). This appraisal tool was used because it helps identify the quality of the research broadly and the rationale for the approach. Additionally, it is a helpful tool when considering multiple studies that use different designs. However, on reflection the appraisal tool did not aid considerations about how the data was analysed and could have been more detailed.

Table 1

Mixed Methods Appraisal Tool (MMAT) version 2018 (Hong et al., 2019).

Category of study designs	Methodological quality criteria	Study: Ebbels & van der Lely (2001)	Ebbels (2007)	Ebbels et al. (2007)	Ebbels et al. (2014)	Ebbels et al. (2024)
Screening questions	S1. Are there clear research questions?	Y	Y	Y	Y	Y
	S2. Do the collected data allow to address the research question?	Y	Y	Y	Y	Y
Quantitative randomized controlled trials	2.1. Is randomization appropriately performed?			Y – although small sample size	Y – although small sample size	
	2.2. Are the groups comparable at baseline?			Y	Y	
	2.3. Are there complete outcome data?			Y	Y	

	2.4. Are outcome assessors blinded to the intervention provided?			Y	Y	
	2.5. Did the participants adhere to the assigned intervention?			Y	Y	
Quantitative non-randomized	3.1. Are the participants representative of the target population.	Y – to an extent (have some form of LI but there is disagreement in the literature about criteria for DLD)	Y – to an extent (have some form of LI but there is disagreement in the literature about criteria for DLD)			Y – to an extent (have some form of LI but there is disagreement in the literature about criteria for DLD)
	3.2. Are measurements appropriate regarding both the outcome and intervention (or exposure)?	Y	Y			Y
	3.3. Are there complete outcome data?	Y	Y			Y
	3.4. Are the confounders accounted for in the design and analysis?	Multiple baseline design accounts for confounding variables but specific confounders were not discussed.	Multiple baseline design accounts for confounding variables but specific confounders were not discussed.			Multiple baseline design accounts for confounding variables but specific confounders were not discussed.

3.5. During the study period, is the intervention administered (or exposure occurred) as intended?	Y	Y	Y