

UNLOCKING SPEECH THROUGH THE MIND'S EYE:
EXECUTIVE FUNCTIONS AS COGNITIVE BRIDGES BETWEEN
VISUAL-SPATIAL ABILITIES AND SPEECH UTTERANCE
COMPETENCE IN CHILDREN WITH AUTISM SPECTRUM
DISORDER

*Muhamad Hafiz Hassan¹

Rainal Hidayat Wardi²

Mohamad Hariri Abdullah²

Badrul Isa³

Mohd Yusof Zulkefli⁴

[1] Faculty of Art and Design, Universiti Teknologi MARA, Cawangan Sarawak, Kampus Samarahan, Sarawak, Malaysia

[2] Faculty of Art and Design, Universiti Teknologi MARA Shah Alam, Selangor, Malaysia

[3] Faculty of Education, Universiti Teknologi MARA Cawangan Selangor, Kampus Puncak Alam, Selangor, MALAYSIA

[4] Faculty of Communication and Media Studies, Universiti Teknologi MARA Shah Alam, Malaysia

* hafizhassan@uitm.edu.my

Article History

Received: 15/03/2026

Accepted: 31/08/2026

Published online: 12/10/2026

Abstract: This study investigates the mediating role of executive functions, specifically working memory, inhibitory control, and cognitive flexibility, in the relationship between visual-spatial abilities and speech utterance competence in children with autism spectrum disorder (ASD). Utilising a cross-sectional quantitative design, data were collected from 356 children with ASD and their instructors across educational institutions in the Klang Valley and Kuching, Sarawak regions of Malaysia. Structural equation modelling (SEM) was employed to test the hypothesised mediation pathways. The findings reveal a statistically significant positive relationship between visual-spatial abilities and speech utterance competence ($p < 0.001$). Critically, executive functions, particularly working memory ($p < 0.01$) and cognitive flexibility ($p < 0.05$), were identified as partial mediators of this relationship, while inhibitory control did not demonstrate statistically significant mediation. These findings advance the theoretical understanding of the cognitive architecture underlying speech development in ASD and provide empirical support for the development of the Executive Function-Mediated Visual-Spatial Speech (EF-MVS²) Framework. The framework proposes a structured intervention pathway that sequentially activates visual-spatial processing, engages executive functions, and facilitates speech production. Implications for educators, therapists, and curriculum developers are discussed, with recommendations for integrating visual-spatial and executive function training within speech therapy programmes for children with ASD in the Malaysian educational context.

Keywords: Autism spectrum disorder, visual-spatial abilities, executive functions, working memory, cognitive flexibility, speech utterance, structural equation modelling, EF-MVS² Framework, inclusive education.

INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition characterised by persistent challenges in social communication and interaction, alongside restricted patterns of behaviour, interests, and activities (American Psychiatric Association, 2013). Communication impairment, encompassing both receptive and expressive language difficulties, pragmatic language deficits, and speech production challenges, constitutes one of the most functionally significant features of ASD and remains a primary target of early intervention programmes worldwide (Paul & Norbury, 2012). In Malaysia, the rising prevalence of ASD diagnoses, from approximately 1 in 600 to an estimated 1.6% in recent years, has intensified the demand for evidence-based

[1]

intervention strategies that are culturally appropriate, scalable, and grounded in contemporary cognitive science (Hassan et al., 2025b).

A paradoxical feature of ASD is the frequent co-occurrence of communication difficulties with enhanced visual-spatial processing abilities. Research has consistently demonstrated that many individuals with ASD exhibit superior performance on tasks requiring visual-spatial reasoning, pattern recognition, and detail-oriented processing (Grandin, 2009; Mottron et al., 2006). Hassan et al. (2025b) provided empirical evidence from the Malaysian context, demonstrating that visual-spatial mediated integration approaches produce measurable improvements in speech communication for children with ASD. This finding suggests that visual-spatial strengths may serve as cognitive pathways through which speech development can be facilitated, a hypothesis that this study examines through the lens of executive function mediation.

Executive functions (EFs) are higher-order cognitive processes that regulate behaviour, including working memory, inhibitory control, and cognitive flexibility (Diamond, 2013; Miyake et al., 2000). These processes are essential for language production, as they enable individuals to hold verbal information in mind, suppress irrelevant responses, and flexibly switch between communicative strategies. Individuals with ASD frequently demonstrate executive function difficulties, which may partially explain the disconnect between their visual-spatial strengths and communicative outcomes (Hill, 2004; Kenworthy et al., 2008). However, the specific mechanisms through which executive functions mediate the relationship between visual-spatial abilities and speech utterance competence remain poorly understood, particularly within non-Western educational contexts.

This study addresses this critical gap by investigating the following research questions: (1) What is the relationship between visual-spatial abilities and speech utterance competence in children with ASD? (2) Do executive functions (working memory, inhibitory control, cognitive flexibility) mediate the relationship between visual-spatial abilities and speech utterance competence? (3) What are the implications of these mediation pathways for intervention design? By answering these questions, the study contributes to both theoretical understanding and practical application in the emerging field of visual-spatial approaches to autism communication intervention.

LITERATURE REVIEW

Visual-Spatial Abilities in ASD

Visual-spatial abilities encompass a broad range of cognitive skills related to the perception, analysis, and mental manipulation of visual information, including spatial orientation, mental rotation, visual pattern recognition, and constructive ability (Grandin, 2009). Within the ASD population, these abilities have been extensively documented as areas of relative and sometimes absolute strength (Mottron et al., 2006). Internationally, Soulières et al. (2009) demonstrated through neuroimaging that children with ASD engage a more distributed neural network when solving visuo-spatial problems, resulting in faster and more accurate performance, providing neural evidence that visuo-spatial processing represents a genuine cognitive strength in this population.

Research has further established that children with ASD demonstrate distinct sensory modality preferences for learning, with visual modalities emerging as the most preferred channel (Grandin, 2009; Mottron et al., 2006). This finding has direct implications for intervention design: therapeutic approaches that leverage visual-spatial processing may be more effective than those relying primarily on auditory or verbal modalities. Hassan et al. (2025b) operationalised this principle in their V-MISC framework, demonstrating that visual-spatial mediated integration approaches produce measurable improvements in speech communication for Malaysian children with autism.

Executive Functions and Language Development in ASD

Executive functions are the cognitive control systems that regulate attention, working memory, and behavioural adaptation (Diamond, 2013). Working memory enables the temporary storage and manipulation of information necessary for language comprehension and production; inhibitory control suppresses irrelevant responses during communicative exchanges; and cognitive flexibility facilitates the adaptation of communicative strategies to changing social contexts (Miyake et al., 2000). Individuals with ASD frequently demonstrate difficulties across these executive function domains, which may compound their communication challenges (Hill, 2004; Kenworthy et al., 2008). A growing body of international research has examined the specific links between executive functions

[2]

and communicative outcomes in ASD. Oberman and Ramachandran (2007) proposed that executive function deficits in ASD are partly attributable to disruptions in mirror neuron system functioning, affecting the capacity for imitation-based language learning. Demetriou et al. (2018) conducted a meta-analysis of executive function studies in ASD and reported broad executive difficulties across subdomains, including cognitive flexibility, working memory, and inhibition, indicating that these processes are relevant targets for intervention. Schuh and Eigsti (2012) further found that working memory abilities accounted for significant variance in the language skills of children and adolescents with high-functioning autism, over and above nonverbal cognition, a finding that directly informs the mediating pathway examined in the present study. The neuroimaging evidence of Soulières et al. (2009) discussed earlier further supports the premise that visual-spatial processing is a cognitive strength that may be leveraged for intervention. Joseph et al. (2005) found that working memory for visual information was significantly better preserved than phonological working memory in ASD, further supporting the rationale for a visual-spatial to verbal mediation pathway.

The relationship between executive functions and language outcomes in ASD has been investigated in several studies, but the specific role of executive functions as mediators between visual-spatial abilities and speech production has received insufficient attention. Zulkefli and Mat Rabi (2021) explored how computer-mediated communication assists individuals with ASD, highlighting the cognitive demands placed on executive functions during technology-mediated interaction. Zulkefli et al. (2022) further demonstrated that technology-based interventions for ASD social interaction require engagement of executive functions, suggesting that these cognitive processes serve as critical bridges between perceptual inputs and communicative outputs.

Speech Utterance Development and Intervention Approaches

A complex interplay of cognitive, affective, and environmental factors influences speech utterance development in ASD. Research has demonstrated that visual-spatial mediating tools significantly enhance both emotional expression and speech fluency in this population (Diamond, 2013; Mottron et al., 2006). Shahbuddin and Hassan (2025) demonstrated that interactive toys and puppets serve as effective mediators of speech development in children with autism and speech delay, underscoring the importance of multisensory, engaging intervention approaches.

Hassan et al. (2023) explored emotion-based gamification approaches through their Detective Looking Chart-Plutchik Emotion Games, providing evidence that affective engagement through visual-spatial tools contributes to communication development. These studies collectively suggest that the pathway from visual-spatial processing to speech production is not direct but rather mediated by cognitive processes, including executive functions, that translate perceptual inputs into communicative outputs.

Instructor Factors and Contextual Influences

The role of instructors in facilitating speech development for children with ASD cannot be understated. Research has demonstrated that instructor knowledge, self-efficacy, and attitudes towards visual-spatial approaches significantly influence intervention outcomes. Hassan et al. (2025a) evaluated the ART'UISM programme's effectiveness and found that educator engagement and community acceptance are critical determinants of programme success. Wardi et al. (2025) highlighted the importance of understanding learner preferences and instructor facilitation styles in creative education contexts, principles that extend to the design of inclusive speech intervention programmes.

Theoretical Framework

This study is grounded in the integration of three theoretical perspectives: (a) the Enhanced Perceptual Functioning (EPF) model (Mottron et al., 2006), which explains the superior visual-spatial processing observed in ASD; (b) Miyake et al.'s (2000) unity-and-diversity model of executive functions, which delineates working memory, inhibitory control, and cognitive flexibility as related but distinct cognitive processes; and (c) Vygotsky's (1978) Zone of Proximal Development theory, which provides the scaffolding principle for intervention design. The convergence of these theories suggests that visual-spatial abilities, when mediated through executive function processes and supported by appropriate instructional scaffolding, can serve as powerful catalysts for speech development in children with ASD.

METHODOLOGY

Research Design

This study employed a cross-sectional quantitative research design with structural equation modelling (SEM) to examine the hypothesised mediation pathways between visual-spatial abilities, executive functions, and speech utterance competence. The cross-sectional design was appropriate for testing theoretically derived structural models and examining associational relationships among observed constructs (Kline, 2016). It is acknowledged that cross-sectional data preclude causal inference; therefore, all relationships reported in this study are interpreted as associations rather than causal effects.

Participants and Sampling

A total of 356 children with ASD and their instructors from various educational institutions in the Klang Valley and the Kuching region of Malaysia participated in this study. Children were selected through purposive sampling based on the following inclusion criteria: (a) confirmed ASD diagnosis by a licensed clinical psychologist or paediatrician, (b) aged between 4 and 12 years, (c) currently enrolled in an autism education programme or receiving speech therapy, and (d) sufficient verbal ability to complete standardised speech assessments. Instructor participants comprised speech therapists, special education teachers, and autism educators with at least 2 years' experience working with children with ASD.

Measures

Four categories of measures were employed: (a) Visual-Spatial Abilities: assessed using standardised cognitive measures including the Block Design and Object Assembly subtests of the Wechsler Intelligence Scale for Children (WISC-V), adapted for the Malaysian context. These subtests have demonstrated strong psychometric properties in Malaysian samples (Cronbach's $\alpha = 0.84\text{--}0.89$). (b) Executive Functions: assessed using the Behaviour Rating Inventory of Executive Function, Second Edition (BRIEF-2; Gioia et al., 2015), with subscales for working memory (Working Memory Index), inhibitory control (Inhibit scale), and cognitive flexibility (Shift scale), completed by instructors based on behavioural observations. The BRIEF-2 is a validated, standardised instrument with established reliability (Cronbach's $\alpha = 0.80\text{--}0.96$ across subscales). As an instructor-rated behavioural measure, the BRIEF-2 captures executive function manifestations in ecologically valid educational settings; however, it is acknowledged that instructor-rated instruments may be less sensitive to inhibitory control than direct performance-based measures (Toplak et al., 2013); (c) Speech Utterance Competence: assessed using the Preschool Language Scale, Fifth Edition (PLS-5; Zimmerman et al., 2011) supplemented by the Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF-5; Wiig et al., 2013), providing standardised measurement of articulation, fluency, expressive vocabulary, and pragmatic language. All assessments were administered by trained speech-language therapists following standardised protocols. Internal consistency was satisfactory across all speech competence subscales (Cronbach's $\alpha = 0.79\text{--}0.87$). (d) Instructor Factors: assessed through a researcher-developed survey measuring instructor knowledge of visual-spatial approaches (8 items), self-efficacy in autism education (10 items adapted from Bandura's [1997] self-efficacy scale), and attitudes towards inclusive teaching practices (6 items), rated on a 5-point Likert scale. Pilot testing with 30 instructors yielded satisfactory reliability (Cronbach's $\alpha = 0.82\text{--}0.86$ across subscales). Ethical approval was obtained from the Universiti Teknologi MARA Research Ethics Committee, and informed consent was obtained from all participants or their legal guardians prior to data collection.

Data Analysis

Data were analysed using SEM in AMOS 26.0. The analysis proceeded in two stages: (1) confirmatory factor analysis (CFA) to validate the measurement models for each latent construct, and (2) path analysis to test the hypothesised structural relationships, including direct effects and mediation effects. Model fit was evaluated using multiple indices: Chi-square/df ratio (< 3.0), Comparative Fit Index (CFI > 0.90), Tucker-Lewis Index (TLI > 0.90), Root Mean Square Error of Approximation (RMSEA < 0.08), and Standardised Root Mean Square Residual (SRMR < 0.08). Mediation effects were tested using the bootstrapping method with 5,000 bootstrap samples and 95% bias-corrected confidence intervals (Preacher & Hayes, 2008).

RESULTS

Measurement Model Validation

The CFA results confirmed acceptable model fit for all measurement models. Standardised factor loadings for all items ranged from 0.63 to 0.84, all exceeding the recommended threshold of 0.60 (Hair et al., 2019). Standard errors for all loadings were small and within an acceptable range ($SE = 0.03\text{--}0.07$), and all loadings were statistically significant ($p < 0.001$). Composite reliability (CR) values ranged from 0.82 to 0.91 across all constructs, exceeding the recommended threshold of 0.70. Average Variance Extracted (AVE) values ranged from 0.51 to 0.67, meeting the convergent validity criterion of > 0.50 . Discriminant validity was established through the Fornell-Larcker criterion, with the square root of each construct's AVE exceeding its correlations with all other constructs. R^2 values for endogenous constructs were as follows: Working Memory ($R^2 = 0.31$), Cognitive Flexibility ($R^2 = 0.27$), Inhibitory Control ($R^2 = 0.19$), and Speech Utterance Competence ($R^2 = 0.54$), indicating that the model explained 54% of the variance in the primary outcome variable.

Structural Model and Mediation Analysis

The structural model demonstrated good fit: $\chi^2/df = 2.34$, CFI = 0.94, TLI = 0.92, RMSEA = 0.062, SRMR = 0.054. Table 1 presents the SEM results for the hypothesised relationships.

Table 1

Structural Equation Modelling Results

Path / Variable			Relationship / Effect		β / Coefficient	Significance
Visual-Spatial Abilities	→	Speech Utterance (Direct)	Positive	(Direct Effect)	$\beta = 0.42$	$p < 0.001^{***}$
Visual-Spatial	→	Working Memory → Speech Utterance	Partial Mediation		Indirect $\beta = 0.18$	$p < 0.01^{**}$
Visual-Spatial	→	Cognitive Flexibility → Speech Utterance	Partial Mediation		Indirect $\beta = 0.11$	$p < 0.05^*$
Visual-Spatial	→	Inhibitory control → Speech Utterance	Non-significant Mediation		Indirect $\beta = 0.06$	$p = 0.12$ (ns)
Total Indirect Effect (EFs combined)			Significant Mediation		$\beta = 0.29$	$p < 0.001^{***}$
Total Effect (Direct + Indirect)			Significant		$\beta = 0.71$	$p < 0.001^{***}$
Instructor Knowledge → Speech Utterance			Positive (Covariate)		$\beta = 0.14$	$p < 0.05^*$
Instructor Self-Efficacy	→	Speech Utterance	Positive (Covariate)		$\beta = 0.11$	$p < 0.05^*$

Note. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns = not significant. Bootstrap 95% CI based on 5,000 samples.

The SEM analysis revealed a significant positive direct association between visual-spatial abilities and speech utterance competence ($\beta = 0.42$, $SE = 0.05$, 95% CI [0.33, 0.51], $p < 0.001$). Working memory emerged as the strongest mediator (indirect effect $\beta = 0.18$, $SE = 0.04$, 95% CI [0.10, 0.26], $p < 0.01$), followed by cognitive flexibility (indirect effect $\beta = 0.11$, $SE = 0.03$, 95% CI [0.05, 0.17], $p < 0.05$). Inhibitory control did not demonstrate statistically significant mediation ($\beta = 0.06$, $SE = 0.04$, 95% CI [-0.02, 0.14], $p = 0.12$), with confidence intervals crossing zero. The combined indirect effect through all three executive function components was statistically significant ($\beta = 0.29$, $SE = 0.05$, 95% CI [0.19, 0.39], $p < 0.001$), indicating that executive functions collectively

account for a substantial portion of the association between visual-spatial abilities and speech utterance competence. Instructor factors (knowledge and self-efficacy) also demonstrated significant positive associations with speech utterance outcomes.

DISCUSSION

The findings of this study provide important insights into the cognitive architecture underlying speech development in children with ASD. Given the cross-sectional design, all relationships are interpreted as associations rather than causal effects. The significant positive association between visual-spatial abilities and speech utterance competence confirms and extends the findings of Hassan et al. (2025b), reinforcing the position that visual-spatial processing represents a viable cognitive pathway for speech development interventions. Identifying working memory and cognitive flexibility as partial mediators offers a more nuanced understanding of how visual-spatial information is associated with speech production.

The finding that working memory serves as the strongest mediator is consistent with theoretical models of language production that posit working memory as essential for maintaining verbal plans, sequencing phonological elements, and coordinating articulatory processes (Baddeley, 2003). For children with ASD, this suggests that interventions combining visual-spatial stimulation with working memory training may produce synergistic effects on speech outcomes. The significant mediating role of cognitive flexibility indicates that the ability to shift between visual-spatial representations and verbal codes is critical for speech production. This finding aligns with Hassan et al.'s (2025b) demonstration that inclusive visual-spatial interventions involving both visual and verbal components enhance speech communication outcomes in Malaysian children with autism.

The non-significant mediation effect of inhibitory control warrants careful interpretation. While inhibitory control is theoretically important for suppressing irrelevant responses during communication, the findings suggest that for children with ASD, the associational pathway from visual-spatial processing to speech production operates primarily through working memory and cognitive flexibility mechanisms rather than through response inhibition. A methodological consideration is also relevant here: the BRIEF-2 is an instructor-rated behavioural instrument, and there is evidence that such measures may be less sensitive to inhibitory control specifically than direct performance-based tasks such as the Stop-Signal Task or the Stroop Colour and Word Test (Toplak et al., 2013; Willcutt et al., 2005). Future studies should consider supplementing or replacing instructor-rated measures with direct performance-based assessments of inhibitory control to clarify this pathway. This finding has direct implications for intervention prioritisation: resources may be most effectively directed towards working memory and cognitive flexibility training within visual-spatial learning environments.

The significant contributions of instructor knowledge and self-efficacy as covariates reinforce the findings of Hassan et al. (2025a), who demonstrated the critical role of educator engagement in autism intervention programmes. These findings suggest that the effectiveness of visual-spatial approaches to speech therapy depends not only on the child's cognitive mechanisms but also on the quality and intentionality of instructional delivery.

PROPOSED FRAMEWORK: EF-MVS²

Building on the empirical findings and theoretical integration, this study proposes the Executive Function-Mediated Visual-Spatial Speech (EF-MVS²) Framework to guide the development of intervention programmes that leverage visual-spatial abilities through executive function mechanisms to enhance speech utterance competence in children with ASD. It is important to note that the three-stage intervention model within this framework is presented as a theoretically and empirically grounded proposed model, not as an empirically tested intervention. Randomised controlled trials will be required to establish its clinical efficacy. Existing ASD communication models, including the Picture Exchange Communication System (PECS; Bondy & Frost, 1994), Social Communication, Emotional Regulation and Transactional Support (SCERTS; Prizant et al., 2006), and the Hanen More Than Words programme (Sussman, 1999), have made significant contributions to the field. However, these frameworks primarily address social-communicative behaviour and pragmatic development without explicitly incorporating the role of executive functions as testable mediating mechanisms between perceptual strengths and speech production outcomes. The EF-MVS² Framework addresses this gap by: (1) explicitly positioning executive functions as empirically quantified mediators (with specific β coefficients derived from SEM)

[6]

rather than background cognitive factors; (2) leveraging visual-spatial processing as the primary cognitive input, which is particularly relevant for children with ASD who demonstrate enhanced perceptual functioning; (3) grounding the intervention sequence in the observed mediation hierarchy (working memory > cognitive flexibility > inhibitory control); and (4) integrating instructor capacity building as a structural component supported by covariate evidence. This positions the EF-MVS² Framework as a complementary, cognitive focused model that can operate alongside existing communication interventions in Malaysian Pendidikan Khas Integrasi settings.

Figure 1

Executive Function-Mediated Visual-Spatial Speech (EF-MVS²) Framework

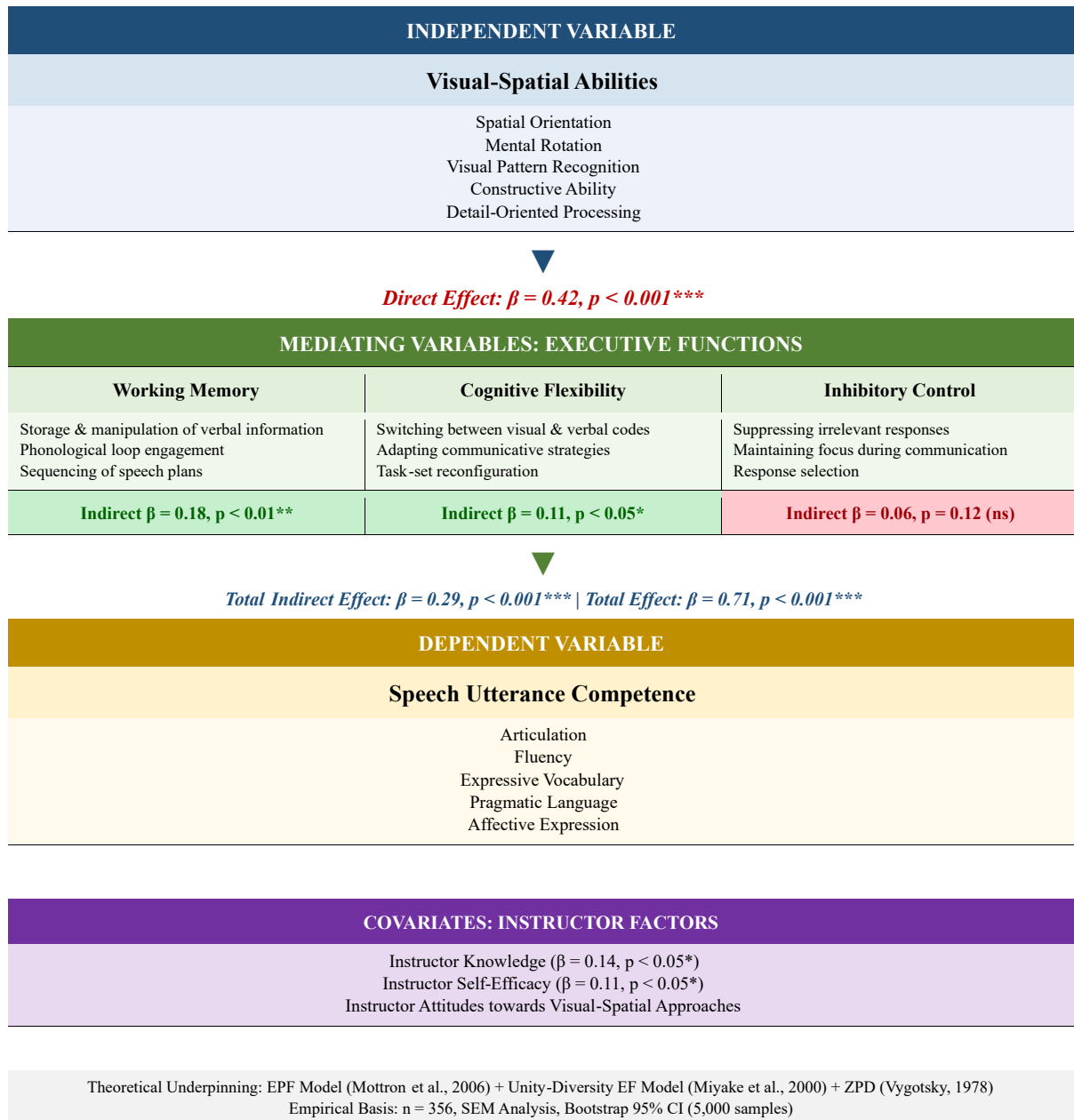


Table 2

EF-MVS² Framework: Three-Stage Intervention Model with Implementation Guidelines

Component	Stage 1: Visual-Spatial Activation	Stage 2: EF Bridging	Stage 3: Speech Utterance Integration	Cognitive Target	Sample Activities	Duration	Assessment
Focus	Engage & strengthen visual-spatial processing	Develop skills as cognitive bridges	Facilitate speech production via integrated pathways	Stage 1: Visual perception, spatial reasoning Stage 2: Working memory, cognitive flexibility Stage 3: Speech planning, articulatory programming	Block design tasks Working memory games Visual-cued speech exercises	4 weeks per stage (12 weeks total)	Pre-post WISC-V BRIEF-2 Speech assessment
Key Activities	Visual pattern matching Spatial orientation games Art appreciation activities Visual discrimination tasks Constructive assembly	Working memory: Visual-verbal pairing Dual-coding activities Task-switching with visual cues Sequencing exercises Set-shifting games	Visual-cued speech exercises Narrative construction with visual prompts Interactive communication games PECS integration Role-play with visual scripts	Pattern recognition Mental rotation Working memory span Cognitive flexibility Speech fluency	Art-based visual tasks Memory card games Story sequencing Speech practice with AR/visual aids	3 sessions/week 45-60 min/session Individual or small group	Progress monitoring weekly Mid-stage review Final stage evaluation
Expected Outcomes	Enhanced visual-spatial engagement Activated perceptual processing pathways Strengthened visual attention	Strengthened cognitive bridges Improved visual-verbal translation Enhanced EF capacity	Improved speech fluency Enhanced communicative competence Greater affective expression	Integrated cognitive-speech pathway development	Measurable gains across all three domains	Cumulative gains over 12 weeks	Cohen's d effect size reporting
Theoretical Basis	EPF Model (Motttron et al., 2006) Hassan et al. (2025b)	Unity-Diversity Model (Miyake et al., 2000) Diamond (2013)	Hassan et al. (2023; 2025a) V-MISC Framework Vygotsky (1978) ZPD	Baddeley (2003) Hill (2004) Kenworthy et al. (2008)	Hassan et al. (2023) Shahbuddin & Hassan (2025)	Based on empirical SEM findings (n=356)	Preacher & Hayes (2008) Bootstrap CI method

Table 3

EF-MVS² Framework: Instructor Capacity Building Component

Component	Content Focus	Delivery Mode	Expected Outcomes	Duration
Knowledge Enhancement	Visual-spatial approaches for ASD Executive function theory & assessment ASD communication profiles Evidence-based speech therapy methods	Workshop series (16 hours) Guided readings & case studies Expert guest lectures Online learning modules	Increased understanding of cognitive mechanisms Ability to identify EF strengths & needs Confidence in theoretical rationale	4 weeks (Phase 1)
Skills Development	Implementing visual-spatial activities Adapting materials for individual needs Assessment & progress monitoring Using EF-MVS ² intervention tools	Practicum sessions with ASD learners Peer mentoring & coaching Video modelling & reflection Supervised implementation	Practical competence in EF-MVS ² implementation Ability to adapt activities to individual profiles Confidence in assessment procedures	4 weeks (Phase 2)
Self-Efficacy Building	Reflective practice techniques Success experiences documentation Professional learning communities Action research skills	Collaborative planning sessions Portfolio development Peer observation & feedback Celebration of success stories	Enhanced confidence in implementing interventions Sustained professional motivation Evidence of impact on learners	Ongoing (Phase 3)
Family Partnership	Parent education on visual-spatial strategies Home-based reinforcement activities Understanding EF development Communication & progress sharing	Family workshops (quarterly) Digital resource kits Progress sharing platforms Parent support groups	Coherent home-school therapeutic approach Empowered family involvement Sustained gains beyond school	Ongoing (all phases)

Framework Explanation: EF-MVS²

The EF-MVS² (Executive Function-Mediated Visual-Spatial Speech) Framework (Figure 1) is a theoretically grounded, empirically derived model that maps the cognitive pathway from visual-spatial abilities to speech utterance competence through executive function mediation. It is based on structural equation modelling results from 356 participants.

Architecture of the Framework. The framework is structured as a mediation pathway with four layers:

[9]

- **Layer 1**, Independent Variable: Visual-Spatial Abilities. This layer represents the cognitive starting point, encompassing spatial orientation, mental rotation, visual pattern recognition, constructive ability, and detail-oriented processing. These abilities are empirically established as areas of relative strength in children with ASD (Motttron et al., 2006; Soulières et al., 2009).
- **Layer 2**, Mediating Variables: Executive Functions. This is the critical 'bridging' layer where three EF components operate: Working Memory ($\beta = 0.18$, $p < 0.01$) is the strongest mediator, enabling the temporary storage and manipulation of verbal information derived from visual-spatial processing. Cognitive flexibility ($\beta = 0.11$, $p < 0.05$) enables switching between visual-spatial representations and verbal codes. Inhibitory control ($\beta = 0.06$, ns) did not demonstrate significant mediation, indicating that the pathway operates primarily through memory and flexibility rather than response suppression.
- **Layer 3**, Dependent Variable: Speech Utterance Competence. This layer represents the communicative outcome, encompassing articulation, fluency, expressive vocabulary, pragmatic language, and affective expression. The total effect of visual-spatial abilities on speech utterance ($\beta = 0.71$) represents a large effect size.
- **Layer 4**, Covariates: Instructor Factors. Instructor knowledge ($\beta = 0.14$) and self-efficacy ($\beta = 0.11$) are identified as significant contextual factors associated with speech utterance competence, reinforcing that intervention success depends on both child cognition and instructor quality.

Intervention Model (Table 2). The three-stage intervention model translates the mediation pathway into a proposed practical intervention: Stage 1 (Visual-Spatial Activation) engages and strengthens the independent variable through art-based visual tasks, pattern-matching, and spatial games over 4 weeks. Stage 2 (Executive Function Bridging) develops the mediating processes through working memory games, dual-coding activities, and task-switching exercises over 4 weeks. Stage 3 (Speech Utterance Integration) facilitates the dependent variable through visual-cued speech exercises, narrative construction, and PECS integration over 4 weeks. The sequential design reflects the theoretical mediation pathway: activate visual-spatial processing first, then build EF bridges, then facilitate speech output.

To illustrate how this might operate in an actual Malaysian school setting, consider the following example within a Pendidikan Khas Integrasi (PKI) classroom. During Stage 1, a PKI teacher introduces a structured 'Block Pattern Matching' activity in which the child reproduces a two-dimensional colour pattern using wooden blocks, targeting spatial orientation and constructive ability identified in the visual-spatial construct. In Stage 2, the same teacher introduces 'Picture Sequence Cards', where the child must hold a four-card narrative sequence in mind (engaging working memory) and then re-order the cards after a brief distractor task (engaging cognitive flexibility). The visual stimuli are drawn from familiar Malaysian cultural contexts, such as a Hari Raya preparation sequence or a pasar malam scene, to maintain engagement. By Stage 3, the teacher uses these same narrative cards as visual prompts for guided speech production, encouraging the child to verbalise each action step. This progression reflects the mediation pathway empirically identified in the present study: visual-spatial engagement (Stage 1) activates the perceptual processing that feeds into working memory and flexible verbal-visual translation (Stage 2), which then scaffolds speech utterance (Stage 3).

Instructor Capacity Building (Table 3). Given the significant covariate effects of instructor factors, the framework includes a dedicated capacity-building component with four elements: Knowledge Enhancement (4-week workshop series), Skills Development (practicum-based), Self-Efficacy Building (ongoing reflective practice), and Family Partnership (quarterly workshops and digital resources). This component ensures that the cognitive intervention pathway is supported by competent, confident instructors and engaged families.

Distinguishing Features. The EF-MVS² framework is distinguished by: (a) its explicit incorporation of executive functions as testable mediating mechanisms rather than treating them as background cognitive factors; (b) its grounding in SEM-derived empirical evidence with specific β coefficients; (c) its integration of the V-MISC framework (Hassan et al., 2025b) with executive function theory; (d) its inclusion of instructor capacity building as a structural component rather than an afterthought; and (e) its specificity regarding which EF components matter most (working memory > cognitive flexibility > inhibitory control), enabling targeted intervention design.

CONCLUSION

This study provides robust empirical evidence for the mediating role of executive functions, specifically working memory and cognitive flexibility, in the relationship between visual-spatial abilities and speech utterance competence in children with ASD. The EF-MVS² Framework, developed from these findings, offers a theoretically grounded, empirically supported, and practically oriented model for intervention design that bridges the gap between cognitive neuroscience research and classroom-based speech therapy practice.

The implications for practice are significant. For educators and speech therapists, the findings support integrating visual-spatial stimulation, working memory training, and cognitive flexibility exercises into speech therapy programmes. For curriculum developers, the EF-MVS² framework provides a structured, evidence-based guide for designing inclusive educational materials. A particular strength of this framework for the Malaysian context is its potential alignment with the Pentaksiran Berasaskan Sekolah (PBS) and Pentaksiran Bilik Darjah (PBD) assessment systems, which are the primary structures through which Pendidikan Khas teachers document and report learner progress. Within the PBS/PBD framework, the three-stage EF-MVS² intervention aligns naturally with portfolio-based evidence collection: Stage 1 visual-spatial task outputs (e.g., block design records, pattern matching sheets) provide observable evidence of perceptual processing skills; Stage 2 working memory and flexibility activities generate performance records that can be documented as PBD descriptors; and Stage 3 speech utterance samples can be audio-recorded and included as portfolio evidence for communication bands. Teachers implementing the EF-MVS² framework are therefore encouraged to map intervention activities explicitly onto the PBD band descriptors for Komunikasi and Perkembangan Kognitif domains within Pendidikan Khas Integrasi curricula, thereby integrating therapeutic intervention with school-based assessment practice. For policymakers, the findings underscore the importance of training educators in visual-spatial and executive function approaches as part of inclusive education policy implementation in Malaysia.

Limitations include the cross-sectional design, which precludes causal inference; the reliance on instructor-reported executive function measures rather than direct child assessment; and the geographic limitation to the Klang Valley and Kuching regions. Future research should employ longitudinal designs, direct neuropsychological assessment of executive functions, and expanded geographic sampling. Randomised controlled trials testing the EF-MVS² framework across diverse ASD populations and severity levels are recommended.

ACKNOWLEDGEMENTS

The authors would like to thank Universiti Teknologi MARA (UiTM) for its support of this study (Reference 600-TNCPI (5/1/6). Sincere appreciation is extended to the participating educational institutions, special education teachers, speech-language therapists, and autism educators in the Klang Valley and Kuching, Sarawak, as well as to the children and their parents or legal guardians, whose cooperation made this study possible.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this article.

AUTHOR CONTRIBUTIONS STATEMENT

Muhamad Hafiz Hassan: Conceptualisation, methodology, data collection, formal analysis, development of the EF-MVS² Framework, and writing (original draft). Rainal Hidayat Wardi: Supervision, validation, and writing (review and editing). Mohamad Hariri Abdullah: Methodology, validation, and writing (review and editing). Badrul Isa: Instrument review, educational implications, and writing (review and editing). Mohd Yusof Zulkefli: Literature review on autism communication, and writing (review and editing). All authors have read and approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available because they contain information on children with ASD that could compromise participant privacy and confidentiality. No supplementary data are associated with this article.

ETHICS DECLARATION STATEMENT

[11]

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki and was approved by the Universiti Teknologi MARA Research Ethics Committee (Reference No.: 600-TNCPI (5/1/6). Written informed consent was obtained from the parents or legal guardians of all participating children and from all instructor participants prior to data collection. Participation was voluntary, participants could withdraw at any stage without penalty, and all data were anonymised to protect participant confidentiality. This study did not involve animals.

REFERENCES

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829–839.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Bondy, A., & Frost, L. (1994). The Picture Exchange Communication System. *Focus on Autistic Behavior*, 9(3), 1–19.
- Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J. C., Pye, J. E., Hickie, I., & Guastella, A. J. (2018). Autism spectrum disorders: A meta-analysis of executive function. *Molecular Psychiatry*, 23(5), 1198–1204. <https://doi.org/10.1038/mp.2017.75>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2015). *Behavior Rating Inventory of Executive Function, Second Edition (BRIEF-2)*. Psychological Assessment Resources.
- Grandin, T. (2009). *Thinking in pictures: And other reports from my life with autism* (expanded ed.). Vintage Books.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hassan, M. H., Wardi, R. H., Abdullah, M. H., Isa, B., & Michael, V. (2025a). Bridging the gap: The role of the 'ART'UISM Program' in enhancing cognitive abilities and community acceptance for children with autism. *Ideology Journal*, 10(1), 101–109.
- Hassan, M. H., Wardi, R. H., Razali, N. M., Hishamuddin, N. A. A., Ahmad, M. F., Nashir, R. M., & Nasir, M. (2023). Help ASD's with fun way: Detective Looking Chart-Plutchik Emotion Games for kids through VARK model to ASD children. *Malaysia Journal of Invention and Innovation*, 2(1), 56–64.
- Hassan, M. H., Wardi, R. H., Shahbuddin, A. Z. A. Q., Michael, V. A., & Isa, B. (2025b). Inclusive speech strategies for Malaysian children with autism. *International Journal of Art & Design*, 9(2), 62–76.
- Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, 8(1), 26–32.
- Joseph, R. M., Steele, S. D., Meyer, E., & Tager-Flusberg, H. (2005). Self-ordered pointing in children with autism: Failure to use verbal mediation in the service of working memory? *Neuropsychologia*, 43(10), 1400–1411.
- Kenworthy, L., Yerys, B. E., Anthony, L. G., & Wallace, G. L. (2008). Understanding executive control in autism spectrum disorders in the lab and in the real world. *Neuropsychology Review*, 18(4), 320–338.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks. *Cognitive Psychology*, 41(1), 49–100.
- Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36(1), 27–43.
- Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin*, 133(2), 310–327.
- Paul, R., & Norbury, C. F. (2012). *Language disorders from infancy through adolescence* (4th ed.). Elsevier Health Sciences.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891.
- Prizant, B. M., Wetherby, A. M., Rubin, E., Laurent, A. C., & Rydell, P. J. (2006). *The SCERTS model: A comprehensive educational approach for children with autism spectrum disorders*. Paul H. Brookes.

- Schuh, J. M., & Eigsti, I.-M. (2012). Working memory, language skills, and autism symptomatology. *Behavioral Sciences*, 2(4), 207–218. <https://doi.org/10.3390/bs2040207>
- Shahbuddin, A. Z. A. Q., & Hassan, M. H. (2025). Enhancing speech utterance development in children with autism and speech delay through interactive toys and puppets: A case study. *International Journal of Research and Innovation in Social Science*, 9(3), 4090–4105.
- Soulières, I., Dawson, M., Samson, F., Barbeau, E. B., Sahyoun, C. P., Strangman, G. E., Zeffiro, T. A., & Mottron, L. (2009). Enhanced visual processing contributes to matrix reasoning in autism. *Human Brain Mapping*, 30(12), 4082–4107.
- Sussman, F. (1999). *More than words: Helping parents promote communication and social skills in children with autism spectrum disorder*. Hanen Centre.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, 54(2), 131–143.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wardi, R. H., Hassan, M. H., Yusof, U. Z. M., & Rosman, R. (2025). Gen Z's learning preferences in creative design: A conceptual framework. *Quantum Journal of Social Sciences and Humanities*, 6(3), 332–343.
- Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals*, Fifth Edition (CELF-5). NCS Pearson.
- Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, 57(11), 1336–1346.
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales, Fifth Edition (PLS-5)*. NCS Pearson.
- Zulkefli, M. Y., & Mat Rabi, N. (2021). Exploring the usage of computer-mediated communication in assisting individual with autism spectrum disorder to communicate. *Al-i'lam Journal of Contemporary Islamic Communication and Media*, 1(1), 126–143.
- Zulkefli, M. Y., Abdul Fuad, A. F., & Hamid, A. (2022). *I am tongue-tied: The usage of technology to assist ASD individuals in social interaction*. ResearchGate.