Perspectives on heterogeneity-informed cognitive training for attention-deficit/hyperactivity disorder

Da-Wei Zhang[1,2]*

[1]Department of Psychology and Center for Place-Based Education, Yangzhou University, Yangzhou, China,
[2]Department of Psychology, Monash University Malaysia, Bandar Sunway, Malaysia

Attention-deficit/hyperactivity disorder (AD/HD) is a heterogeneous neurodevelopmental condition, posing a severe threat to quality of life. Pharmacological therapies are the front-line treatment; however, their shortages encourage the development of alternative treatments for AD/HD. One promising method of developing alternative treatments is cognitive training (CT). A CT-based therapy was recently approved by the US Food and Drug Administration. However, due to heterogeneity in AD/HD, a CT protocol is unlikely to provide a one-size-fits-all solution for all patients with AD/HD. Therefore, this article highlights key aspects that need to be considered to further develop CT protocols for AD/HD, regarding training content, timing, suitability, and delivery mode. The perspectives presented here contribute to optimizing CT as an alternative option for treating AD/HD.

KEYWORDS

ADHD, cognitive training, heterogeneity, training protocols, optimization

1. Introduction

Attention-deficit/hyperactivity disorder (AD/HD), a common neurodevelopmental disorder (1), poses a severe threat to quality of life (2). AD/HD is heterogeneous in etiology, with both hereditability and environmental factors contributing to the condition (3). Pharmacological therapies are the front-line treatment (4, 5) but have been criticized for their side effects (5, 6), lack of long-term effect (7), and economic impact (8). Thus, developing alternative treatments for AD/HD is of great importance.

Recently, cognitive training (CT) has become an alternative option for treating AD/HD. CT, a concept analogous to physical training, is a non-pharmacological approach and provides an array of opportunities for developing desired cognitive skills (9). It has been shown that cognitive impairment can contribute to AD/HD symptoms (10, 11), but that it can also be malleable (12, 13). Based on these, CT aims to alleviate AD/HD symptoms by improving impaired cognitive abilities. Two decades have passed since CT was introduced in AD/HD (14). Kollins et al. (15) reported a landmark study in this field, which demonstrated that a 4-week CT protocol, known as AKL-T01, significantly improved inattention in children with AD/HD with a multi-center randomized control trial. This promising result led to AKL-T01 being approved by the US Food and Drug Administration (FDA) for treating inattention in children with AD/HD—the first FDA-approved digital therapeutics in AD/HD (16).

While CT has become a treatment option, heterogeneity in AD/HD can affect the effect of CT and should be considered for further developing CT protocols. AD/HD is well-known for its heterogeneity in etiology and phenotype (3). Even for the same patient, AD/HD symptoms can fluctuate and result in secondary mental disorders over time (3, 17). Due to heterogeneity, patients with AD/HD respond differently to a range of treatment options, for example, pharmacological therapies (18), neurofeedback training (19), and trigeminal nerve stimulation (20). Similarly, a recent study reported that children with AD/HD differed in their response to CT (21). Together, due to the complexity caused by heterogeneity, a CT protocol is unlikely a one-size-fits-all solution for AD/HD. However, the role of heterogeneity has not been adequately considered. Therefore, this article highlights key aspects that need further consideration to optimize CT as a treatment for AD/HD, regarding training content, timing, suitability, and delivery mode.

2. Beyond executive function training

The multifaceted nature of cognitive deficits in AD/HD requires CT to target a variety of cognitive abilities. Multiple cognitive deficits have been reported in AD/HD at the group level, including sustained attention, executive functions (EF), and motivational functions (10, 11). These cognitive functions are mediated by unique brain mechanisms (11, 22) and lead to different aspects of AD/HD symptoms (11). CT should therefore be directed at addressing these cognitive deficits.

However, few CT protocols considered motivational deficits in AD/HD. Motivation can be incorporated in CT for AD/HD in at least two ways. Firstly, CT protocols can be tailored to train impaired motivational processing in AD/HD. The preference for immediate small rewards over delayed large rewards—the delay gratification issue—is a well-recognized motivational deficit in AD/HD (11). However, a glance at review papers reveals that previous CT protocols for AD/HD mainly targeted sustained attention and executive function (EF) but rarely trained this motivational deficit [e.g., (23)]. In typically developed children, there have been successful attempts to delay gratification by reducing undesired thinking (24). Future studies may explore whether this CT protocol designed to delay gratification can benefit those with AD/HD. Secondly, CT protocols can leverage motivation to boost training of sustained attention and EF. The effectiveness of CT requires multiple sessions of intensive training, which can be challenging for AD/HD as evidenced by the high dropout rate in previous studies [e.g., (25, 26)]. Boosting motivation is an immediate solution to this problem. Motivational brain networks interact with the brain networks underlying sustained attention and EF such that motivation can prompt more efforts to sharpen sustained attention and EF (22). Consequently, CT protocols may involve features to boost motivation for facilitating or enhancing training effects. It has been demonstrated that CT protocols incorporating motivational elements (e.g., gamification) produced better training effects in healthy populations (27, 28). Some CT protocols for AD/HD indeed incorporated motivational features, such as gamifying training tasks and providing immediate feedback (29, 30). However, it should be noted that motivational features (e.g., changing game themes and daily prizes) may become distractors instead and negatively impact training improvement in typical children (31). Since AD/HD is more susceptible to distractors,

inappropriate motivational features may have a more significant detrimental impact. Together, while incorporating motivational design into CT is an effective way to facilitate training effects in AD/HD, future research should examine how to optimize motivational features.

Beyond rehabilitation (e.g., training on motivation and EF), cognitive training may contribute to academic development in AD/HD by improving domain-specific academic skills. Schooling can be particularly challenging for AD/HD. Although AD/HD is not a learning disability, those with AD/HD typically experience learning difficulties beginning in their early schooling years (32). Early learning difficulties adversely affect the development of fundamental domain-specific academic skills, such as mathematical facts retrieval (33) and spelling (34), which in turn negatively impacts their long-term academic achievement (35). In typical children, there has been evidence that specific forms of CT can improve domain-specific academic skills. For example, fundamental mathematical and reading skills can be improved by "number line training" (36) and "phonemic awareness training" (37), respectively. Thus, future studies may examine whether cognitive training targeting domain-specific academic skills is superior to the "teaching as usual" condition in improving early academic skills and easing early learning difficulties in children with AD/HD.

Regardless of training purposes, impurity in training content should be considered. Psychometric studies have shown that cognitive task paradigms, particularly those used for EF, involve not only the targeted cognitive ability but also incidental cognitive abilities (38). The color-word Stroop task, for example, is expected to tax inhibition but also involves semantic and sensory processing. The issue is known as task impurity (38). CT is typically developed based on the task paradigms for measuring desired cognitive abilities. Due to task impurity, repetitive CT also increases exposure to non-target cognitive abilities. An unwanted consequence is that trainees primarily focus and improve on non-target cognitive abilities. One possible solution to the issue is to develop CT protocols based on multiple task paradigms that heavily tax the desired abilities. This is based on the rationale that the impurity component is task-specific (non-overlapping between tasks), whereas the desired component is task-general (overlapping between tasks) (38). Hence, CT based on multiple task paradigms minimizes the involvement of non-target abilities across sessions.

3. Toward personalized CT

The cognitive deficits exhibited at the group level (e.g., ADHD vs. controls) do not necessarily indicate that all patients with AD/HD suffer the same deficits. Instead, cognitive deficits vary among those with AD/HD (39). Some patients with AD/HD may have impaired sustained attention, while others may have typical sustained attention but a deficient subcomponent of EF.

The heterogeneous cognitive profiles offer insight into who may benefit most from CT in AD/HD. In a working memory training program, children with AD/HD who had poorer working memory before training (baseline) improved more (21). The result is in line with the pattern displayed from other types of cognitive interventions in children—a poorer baseline predicts greater training improvement (40). These findings suggest that CT in AD/HD should be tailored

to the deficient type, which is particularly important since there may be various CT protocols (e.g., attention-focused CT vs. motivation-focused CT).

Although the baseline-improvement pattern seems intuitive, it is in contrast to what has been found in typically developed populations—a poorer baseline predicts smaller improvement in CT (41, 42). The discrepancy has not been well-explored. One possible explanation is that CT acts on different mechanisms in different populations. A capacity-efficiency model has been developed to explain CT mechanisms (9). For people who have not fully developed their cognitive capacity (e.g., AD/HD), training might act on the capacity mechanism where CT enhances the capacity of overall cognitive resources (43). Poor baselines, in this case, indicate more room for improvement. Thus, the poorer-greater pattern is observed. Conversely, CT may act mainly on optimizing cognitive performance within the existing capacity limit for people with fully developed cognitive abilities (e.g., typical adults). Poor baselines indicate a lack of knowledge and sources to optimize cognitive performance, resulting in the poorer-poorer pattern.

Neuroimaging profiles may provide unique value in identifying who might benefit from CT in AD/HD. After controlling baseline cognitive profiles, baseline brain structural profiles were found to predict additional CT improvement in typically developing children (41). This may also apply to those with AD/HD. Additionally, neuroimaging profiles have been shown to uniquely predict improvements in other interventions (20). Future studies may examine whether combining cognitive profiles with neuroimaging profiles helps identify those with AD/HD who are most likely to benefit from CT.

The question of who might benefit from CT prompts responder analysis but deserves methodological consideration. Selecting the statistical model is one of the issues. Responder analysis typically begins with defining responders and non-responders based on the difference between the pre- post-intervention change. The cut-off point is arbitrary (e.g., the 30% as responders and the bottom 30% as non-responders). Then, the defined responders are characterized by a group comparison (responder vs. non-responders) or regression analysis. However, this method of determining responders may suffer regression toward the mean because of measurement errors [for more discussion and alternative models, please see Castro-Schilo and Grimm (44) and (30, 45)]. The second issue refers to generalizability. Characterizing responders is a prediction problem. Most studies have addressed this issue by using in-sample regression analysis, which primarily reflects correlations between variables within a sample and does not necessarily generalize to other samples (46). There have been recommendations (e.g., k-fold cross-validation and large sample sizes) for ensuring the validity of predictive models (46).

4. Age matters

Childhood is a unique window for conducting CT in AD/HD. Childhood is a sensitive period for developing cognitive skills because of its greater neuroplasticity (47, 48). CT has been shown to lead to larger cognitive development in children (49). In the same vein, the advantage of childhood may also enhance the effectiveness of CT as a treatment for reducing AD/HD symptoms.

The developmental course of cognitive deficits in AD/HD also favors CT in childhood. Developmental cascades refer to the effect that the development of a cognitive function at one time point consequently affects the emergence or growth of other cognitive functions at later time points (50). The effect of developmental cascades can be adverse. In AD/HD, for instance, atypical cognitive development at an early age may eventually lead to various cognitive deficits or secondary mental disorders at an older age (3, 51). Following this developmental perspective, CT should be delivered as early as possible to avoid the adverse effect of developmental cascades.

The question then arises as to how early CT can be performed. A recent study examined the feasibility of CT intervention for infants aged 9–16 months (52). After easing the concern on feasibility in young children, the question is therefore reduced to how early AD/HD can be reliably diagnosed. Although AD/HD is usually diagnosed at school age, genetic progress may enable an earlier diagnosis. Using the genome-wide association study, researchers have identified several genetic loci associated with AD/HD symptoms [e.g., (53)]. The genetic pattern implies the possibility of conducting preemptive CT for individuals genetically predisposed to AD/HD to prevent the detrimental cascades. However, preemptive interventions in AD/HD will also be subject to the ethical issues related to preemptive interventions in other neurodevelopmental disorders, for example, neglecting the effect of "nurture" and miscommunicating early concerns to families [for further discussion, please see Manzini et al. (54)].

The advantage of early CT does not preclude the possibility of CT for adults with AD/HD. As neuroplasticity is a life-long brain property, adults with AD/HD are also expected to benefit from CT. Despite this, few studies have examined the efficacy of CT in adults with AD/HD, and most of them reported negative results (55–59). There are two possible reasons for the discouraging results. The existing studies primarily used CT protocols that were developed for children. This may not be appropriate, given that the dropout rate for adults is greater than that for children (25). The high dropout rate also resulted in insufficient statistical power to detect training effects (58). The second possible reason is that previous CT studies mainly targeted working memory (55–58), while inattention is the dominant symptom in adults with AD/HD (60, 61). Therefore, CT protocols focusing on attention may be more effective in treating adults with AD/HD. Promising outcomes were indeed reported for attention-based CT in adults with AD/HD (59).

5. Not only computerized CT

In terms of delivery, CT can be broadly divided into computerized CT and non-computerized CT (23). Computerized CT delivers training content using digital devices (e.g., computers and smartphones). Training content is typically gamified to increase training adherence, for example, using a gamified Go/Nogo task to train inhibition (29, 30). Non-computerized CT is a broader term, encompassing multiple forms, such as board games based on the Corsi block task to train working memory (23) and physical activity based on the "Simon Says" task to train inhibition (62). Computerized CT offers the advantages of precisely controlling training content (e.g., training duration and difficulty) and requires less manual

guidance, whereas non-computerized CT is more cost-effective and ecological (45).

Due to the difference in delivery methods, CT may target different aspects of the same cognitive construct. Many CT protocols target EF, as described above. Previous studies have reported that EF measured by computerized cognitive tasks is not or is weakly correlated with those measured by questionnaires [e.g., (63)]. In light of these findings, a distinction is made between "cold" and "hot" EF (e.g., those measured by computerized tasks or questionnaires) (64). The former tax more on "pure" EF, whereas the latter reflects the capacity of EF in everyday life where other cognitive processes (e.g., emotion and motivation) are also involved (64). Non-computerized CT has been suggested advantage in improving "hot" EF (65). It is possible that computerized and non-computerized CT protocols target EF from different aspects—"cold" EF or "hot" EF. However, few studies have examined this issue to date. Future studies should systematically compare the two types of CT in terms of their effects on "cold" and "hot" EF. The comparison could optimize the delivery mode of CT in AD/HD since some patients with AD/HD have impaired "cold" EF, while others have impaired "hot" EF (66).

Combining CT with other interventions is another promising delivery mode. One reason for this combination is to improve the transferability of training effects in specific settings. Compared to CT alone, CT combined with supported employment programs has been shown to generate larger training effects in vocational settings in other populations (67). This combined protocol may also be beneficial to AD/HD with poor vocational skills. The transfer effect of CT is often explained by the common element theory (68, 69). In this regard, the combined protocol is superior because it discloses the similarity between skills learned in CT and skills required in specific settings. Secondly, CT can be used in conjunction with interventions that enhance neuroplasticity. This combination is grounded on the points that neuroplasticity is a prerequisite for CT improvement, and interventions that enhance neuroplasticity can thus improve training effects (45). Following this perspective, attempts have been made to combine CT with non-invasive brain stimulation; however, the results are inconsistent [see Dakwar-Kawar et al. (70) but also Westwood et al. (71)]. The disparity may be due to the different timing and intensity at which brain stimulation was delivered, as the impact of the same brain stimulation on neuroplasticity varies with both time and intensity (45, 72). Further exploration of the mechanisms underlying non-invasive brain stimulation may lead to more effective combinations.

6. Conclusion

Medicines are usually accompanied by a lengthy guide that describes the factors affecting their effectiveness. While FDA has approved one CT protocol as a treatment for AD/HD, the key message from this article is that heterogeneity in AD/HD affects what, who, when, and how CT should be performed. The complexity of AD/HD suggests that one CT protocol is unlikely to provide a one-size-fits-all solution for AD/HD, thus appealing to "heterogeneous" CT protocols.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding author.

Author contributions

D-WZ participated in the conceptualization, writing, and editing process of this manuscript.

Funding

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Polanczyk GV, Willcutt EG, Salum GA, Kieling C, Rohde LA. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. *Int J Epidemiol.* (2014) 43:434–42. doi: 10.1093/ije/dyt261

2. Biederman J, Petty C, Fried R, Fontanella J, Doyle AE, Seidman LJ, et al. Impact of psychometrically defined deficits of executive functioning in adults with attention deficit hyperactivity disorder. *Am J Psychiatry.* (2006) 163:1730–8. doi: 10.1176/ajp.2006.163.10.1730

3. Nigg JT, Sibley MH, Thapar A, Karalunas SL. Development of ADHD: etiology, heterogeneity, and early life course. *Annu Rev Dev Psychol.* (2020) 2:559–83. doi: 10.1146/annurev-devpsych-060320-093413

4. Subcommittee on Attention-Deficit/Hyperactivity Disorder; Steering Committee on Quality Improvement and Management; Wolraich M, Brown L, Brown RT, DuPaul G, Earls M, et al. ADHD: clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics.* (2011) 128:1007-22. doi: 10.1542/peds.2011-2654

5. Caye A, Swanson JM, Coghill D, Rohde LA. Treatment strategies for ADHD: An evidence-based guide to select optimal treatment. *Mol Psychiatry.* (2019) 24:390–408. doi: 10.1038/s41380-018-0116-3

6. Graham J, Banaschewski T, Buitelaar J, Coghill D, Danckaerts M, Dittmann RW, et al. European guidelines on managing adverse effects of medication for

ADHD. *Eur Child Adolesc Psychiatry*. (2011) 20:17–37. doi: 10.1007/s00787-010-0140-6

7. Wang GJ, Volkow ND, Wigal T, Kollins SH, Newcorn JH, Telang F, et al. Long-term stimulant treatment affects brain dopamine transporter level in patients with attention deficit hyperactive disorder. *PLoS ONE*. (2013) 8:e63023. doi: 10.1371/journal.pone.0063023

8. Chorozoglou M, Smith E, Koerting J, Thompson MJ, Sayal K, Sonuga-Barke EJ. Preschool hyperactivity is associated with long-term economic burden: evidence from a longitudinal health economic analysis of costs incurred across childhood, adolescence and young adulthood. *J Child Psychol Psychiatry*. (2015) 56:966–75. doi: 10.1111/jcpp.12437

9. von Bastian CC, Belleville S, Udale RC, Reinhartz A, Essounni M, Strobach T. Mechanisms underlying training-induced cognitive change. *Nat Rev Psychol*. (2022) 1:30–41. doi: 10.1038/s44159-021-00001-3

10. Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol Bull*. (1997) 121:65–94. doi: 10.1037/0033-2909.121.1.65

11. Sonuga-Barke EJ. Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biol Psychiatry*. (2005) 57:1231–8. doi: 10.1016/j.biopsych.2004.09.008

12. Merzenich MM, Van Vleet TM, Nahum M. Brain plasticity-based therapeutics. *Front Hum Neurosci*. (2014) 8:385. doi: 10.3389/fnhum.2014.00385

13. Johnson MH. Interactive specialization: a domain-general framework for human functional brain development? *Dev Cogn Neurosci*. (2011) 1:7–21. doi: 10.1016/j.dcn.2010.07.003

14. Klingberg T, Forssberg H, Westerberg H. Training of working memory in children with ADHD. *J Clin Exp Neuropsychol*. (2002) 24:781–91. doi: 10.1076/jcen.24.6.781.8395

15. Kollins SH, DeLoss DJ, Cañadas E, Lutz J, Findling RL, Keefe RSE, et al. A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): a randomised controlled trial. *Lancet Digit Health*. (2020) 2:e168-e178. doi: 10.1016/S2589-7500(20)30017-0

16. FDA. *FDA Permits Marketing of First Game-Based Digital Therapeutic to Improve Attention Function in Children with ADHD* (2020).

17. Sibley MH, Arnold LE, Swanson JM, Hechtman LT, Kennedy TM, Owens E, et al. Variable patterns of remission from ADHD in the multimodal treatment study of ADHD. *Am J Psychiatry*. (2022) 179:142–51. doi: 10.1176/appi.ajp.2021.21010032

18. Cortese S. Pharmacologic treatment of attention deficit–hyperactivity disorder. *N Engl J Med*. (2020) 383:1050–6. doi: 10.1056/NEJMra1917069

19. Voetterl H, van Wingen G, Michelini G, Griffiths KR, Gordon E, DeBeus R, et al. Brainmarker-I differentially predicts remission to various attention-deficit/hyperactivity disorder treatments: a discovery, transfer, and blinded validation study. *Biol Psychiatry Cogn Neurosci Neuroimaging*. (2022) S2451902222000465. doi: 10.1016/j.bpsc.2022.02.007

20. Loo SK, Salgari GC, Ellis A, Cowen J, Dillon A, McGough JJ. Trigeminal nerve stimulation for attention-deficit/hyperactivity disorder: cognitive and electroencephalographic predictors of treatment response. *J Am Acad Child Adolesc Psychiatry*. (2021) 60:856–64.e1. doi: 10.1016/j.jaac.2020.09.021

21. van der Donk MLA, Hiemstra-Beernink A-C, Tjeenk-Kalff AC, van der Leij A, Lindauer RJL. Predictors and moderators of treatment outcome in cognitive training for children with ADHD. *J Atten Disord*. (2020) 24:1914–27. doi: 10.1177/1087054716632876

22. Pessoa L. How do emotion and motivation direct executive control? *Trends Cogn Sci*. (2009) 13:160–6. doi: 10.1016/j.tics.2009.01.006

23. Veloso A, Vicente SG, Filipe MG. Effectiveness of cognitive training for school-aged children and adolescents with attention deficit/hyperactivity disorder: a systematic review. *Front Psychol*. (2020) 10:2983. doi: 10.3389/fpsyg.2019.02983

24. Murray J, Scott H, Connolly C, Wells A. The attention training technique improves children's ability to delay gratification: a controlled comparison with progressive relaxation. *Behav Res Ther*. (2018) 104:1–6. doi: 10.1016/j.brat.2018.02.003

25. Marcelle ET, Ho EJ, Kaplan MS, Adler LA, Castellanos FX, Milham MP. Cogmed working memory training presents unique implementation challenges in adults with ADHD. *Front Psychiatry*. (2018) 9:388. doi: 10.3389/fpsyt.2018.00388

26. Luo X, Guo X, Zhao Q, Zhu Y, Chen Y, Zhang D, et al. A randomized controlled study of remote computerized cognitive, neurofeedback, and combined training in the treatment of children with attention-deficit/hyperactivity disorder. *Eur Child Adolesc Psychiatry*. (2022) 1–12. doi: 10.1007/s00787-022-01956-1 [Epub ahead of print].

27. Deveau J, Lovcik G, Seitz AR. Broad-based visual benefits from training with an integrated perceptual-learning video game. *Vision Res*. (2014) 99:134–40. doi: 10.1016/j.visres.2013.12.015

28. Mohammed S, Flores L, Deveau J, Cohen Hoffing R, Phung C, M. Parlett C, et al. The benefits and challenges of implementing motivational features to boost cognitive training outcome. *J Cogn Enhanc*. (2017) 1:491–507. doi: 10.1007/s41465-017-0047-y

29. Johnstone SJ, Roodenrys SJ, Johnson K, Bonfield R, Bennett SJ. Game-based combined cognitive and neurofeedback training using Focus Pocus reduces symptom severity in children with diagnosed AD/HD and subclinical AD/HD. *Int J Psychophysiol*. (2017) 116:32–44. doi: 10.1016/j.ijpsycho.2017.02.015

30. Zhang D-W, Johnstone SJ, Li H, Luo X, Sun L. Comparing the transfer effects of three neurocognitive training protocols in children with attention-deficit/hyperactivity disorder: a single-case experimental design. *Behav Change*. (2021) 1–19. doi: 10.1017/bec.2021.26

31. Katz B, Jaeggi S, Buschkuehl M, Stegman A, Shah P. Differential effect of motivational features on training improvements in school-based cognitive training. *Front Hum Neurosci*. (2014) 8:242. doi: 10.3389/fnhum.2014.00242

32. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. American Psychiatric Association (2013) 59–65. doi: 10.1176/appi.books.9780890425596

33. Orbach L, Herzog M, Fritz A. Relation of attention deficit hyperactivity disorder (ADHD) to basic number skills and arithmetic fact retrieval in children. *Res Dev Disabil*. (2020) 103:103697. doi: 10.1016/j.ridd.2020.103697

34. Paloyelis Y, Rijsdijk F, Wood AC, Asherson P, Kuntsi J. The genetic association between ADHD symptoms and reading difficulties: the role of inattentiveness and IQ. *J Abnorm Child Psychol*. (2010) 38:1083–95. doi: 10.1007/s10802-010-9429-7

35. Taanila A, Ebeling H, Tiihala M, Kaakinen M, Moilanen I, Hurtig T, et al. Association between childhood specific learning difficulties and school performance in adolescents with and without ADHD symptoms: a 16-year follow-up. *J Atten Disord*. (2014) 18:61–72. doi: 10.1177/1087054712446813

36. Zhang D-W, Zaphf A, Klingberg T. Resting state EEG related to mathematical improvement after spatial training in children. *Front Hum Neurosci*. (2021) 15:698367. doi: 10.3389/fnhum.2021.698367

37. Suggate SP. A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *J Learn Disabil*. (2016) 49:77–96. doi: 10.1177/0022219414528540

38. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn Psychol*. (2000) 41:49–100. doi: 10.1006/cogp.1999.0734

39. Nigg JT, Willcutt EG, Doyle AE, Sonuga-Barke EJS. Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biol Psychiatry*. (2005) 57:1224–30. doi: 10.1016/j.biopsych.2004.08.025

40. Diamond A, Lee K. Interventions shown to aid executive function development in children 4 to 12 years old. *Science*. (2011) 333:959–64. doi: 10.1126/science.1204529

41. Nemmi F, Helander E, Helenius O, Almeida R, Hassler M, Räsänen P, et al. Behavior and neuroimaging at baseline predict individual response to combined mathematical and working memory training in children. *Dev Cogn Neurosci*. (2016) 20:43–51. doi: 10.1016/j.dcn.2016.06.004

42. Wiemers EA, Redick TS, Morrison AB. The influence of individual differences in cognitive ability on working memory training gains. *J Cogn Enhanc*. (2019) 3:174–85. doi: 10.1007/s41465-018-0111-2

43. Zhang D-W, Sauce B. Efficiency and capacity mechanisms can coexist in cognitive training. *Nat Psychol Rev*. (in press). doi: 10.1038/s44159-022-00146-9

44. Castro-Schilo L, Grimm KJ. Using residualized change versus difference scores for longitudinal research. *J Soc Pers Relat*. (2018) 35. doi: 10.1177/0265407517718387

45. Zhang D-W, Johnstone SJ, Sauce B, Arns M, Sun L, Jiang H. Remote neurocognitive interventions for attention-deficit/hyperactivity disorder – opportunities and challenges. *PsyArXiv*. (2022). doi: 10.31234/osf.io/7tnz9

46. Poldrack RA, Huckins G, Varoquaux G. Establishment of best practices for evidence for prediction: a review. *JAMA Psychiatry*. (2020) 77:534. doi: 10.1001/jamapsychiatry.2019.3671

47. Ismail FY, Fatemi A, Johnston MV. Cerebral plasticity: windows of opportunity in the developing brain. *Eur J Paediatr Neurol*. (2017) 21:23–48. doi: 10.1016/j.ejpn.2016.07.007

48. Thompson A, Steinbeis N. Sensitive periods in executive function development. *Curr Opin Behav Sci*. (2020) 36:98–105. doi: 10.1016/j.cobeha.2020.08.001

49. Wass SV, Scerif G, Johnson MH. Training attentional control and working memory – is younger, better? *Dev Rev*. (2012) 32:360–87. doi: 10.1016/j.dr.2012.07.001

50. Masten AS, Cicchetti D. Developmental cascades. *Dev Psychopathol*. (2010) 22:491–5. doi: 10.1017/S0954579410000222

51. Halperin JM, Healey DM. The influences of environmental enrichment, cognitive enhancement, and physical exercise on brain development: can we alter the developmental trajectory of ADHD? *Neurosci Biobehav Rev*. (2011) 35:621–34. doi: 10.1016/j.neubiorev.2010.07.006

52. Goodwin A, Jones EJH, Salomone S, Mason L, Holman R, Begum-Ali J, et al. INTERSTAARS: attention training for infants with elevated likelihood of developing ADHD: a proof-of-concept randomised controlled trial. *Transl Psychiatry*. (2021) 11:644. doi: 10.1038/s41398-021-01698-9

53. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet*. (2019) 51:63–75. doi: 10.1038/s41588-018-0269-7

54. Manzini A, Jones EJH, Charman T, Elsabbagh M, Johnson MH, Singh I. Ethical dimensions of translational developmental neuroscience research in autism. *J Child Psychol Psychiatry.* (2021) 62:1363–73. doi: 10.1111/jcpp.13494

55. Dentz A, Guay M-C, Parent V, Romo L. Working memory training for adults with ADHD. *J Atten Disord.* (2020) 24:918–27. doi: 10.1177/1087054717723987

56. Liu Z-X, Lishak V, Tannock R, Woltering S. Effects of working memory training on neural correlates of Go/Nogo response control in adults with ADHD: a randomized controlled trial. *Neuropsychologia.* (2017) 95:54–72. doi: 10.1016/j.neuropsychologia.2016.11.023

57. Mawjee K, Woltering S, Tannock R. Working Memory training in post-secondary students with ADHD: a randomized controlled study. *PLoS ONE.* (2015) 10:e0137173. doi: 10.1371/journal.pone.0137173

58. Stern A, Malik E, Pollak Y, Bonne O, Maeir A. The efficacy of computerized cognitive training in adults with ADHD: a randomized controlled trial. *J Atten Disord.* (2016) 20:991–1003. doi: 10.1177/1087054714529815

59. White HA, Shah P. Training attention-switching ability in adults with ADHD. *J Atten Disord.* (2006) 10:44–53. doi: 10.1177/1087054705286063

60. Larsson H, Dilshad R, Lichtenstein P, Barker ED. Developmental trajectories of DSM-IV symptoms of attention-deficit/hyperactivity disorder: genetic effects, family risk and associated psychopathology: developmental trajectories of DSM-IV symptoms of ADHD. *J Child Psychol Psychiatry.* (2011) 52:954–63. doi: 10.1111/j.1469-7610.2011.02379.x

61. Pingault J-B, Viding E, Galéra C, Greven CU, Zheng Y, Plomin R, et al. Genetic and environmental influences on the developmental course of attention-deficit/hyperactivity disorder symptoms from childhood to adolescence. *JAMA Psychiatry.* (2015) 72:651. doi: 10.1001/jamapsychiatry.2015.0469

62. Wexler BE, Vitulano LA, Moore C, Katsovich L, Smith SD, Rush C, et al. An integrated program of computer-presented and physical cognitive training exercises for children with attention-deficit/hyperactivity disorder. *Psychol Med.* (2021) 51:1524–35. doi: 10.1017/S0033291720000288

63. Mcauley T, Chen S, Goos L, Schachar R, Crosbie J. Is the behavior rating inventory of executive function more strongly associated with measures of impairment or executive function? *J Int Neuropsychol Soc.* (2010) 16:495–505. doi: 10.1017/S1355617710000093

64. Zelazo PD, Carlson SM. Hot and cool executive function in childhood and adolescence: development and plasticity. *Child Dev Perspect.* (2012) 6:354–60. doi: 10.1111/j.1750-8606.2012.00246.x

65. Moore A, Ledbetter C. The promise of clinician-delivered cognitive training for children diagnosed with ADHD. *J Mental Health Clin Psychol.* (2019) 3:3–8. doi: 10.29245/2578-2959/2019/3.1180

66. Barkley RA. *Executive Functions: What They Are, How They Work, and Why They Evolved.* Guilford Press (2012). p. 897–914.

67. Harvey PD, Depp CA, Rizzo AA, Strauss GP, Spelber D, Carpenter LL, et al. Technology and mental health: state of the art for assessment and treatment. *Am J Psychiatry.* (2022) 190–209. doi: 10.1176/appi.ajp.21121254

68. Gathercole SE, Dunning DL, Holmes J, Norris D. Working memory training involves learning new skills. *J Mem Lang.* (2019) 105:19–42. doi: 10.1016/j.jml.2018.10.003

69. Taatgen NA. The nature and transfer of cognitive skills. *Psychol Rev.* (2013) 120:439–71. doi: 10.1037/a0033138

70. Dakwar-Kawar O, Berger I, Barzilay S, Grossman ES, Cohen Kadosh R, Nahum M. Examining the effect of transcranial electrical stimulation and cognitive training on processing speed in pediatric attention deficit hyperactivity disorder: a pilot study. *Front Hum Neurosci.* (2022) 16:791478. doi: 10.3389/fnhum.2022.791478

71. Westwood SJ, Criaud M, Lam SL, Lukito S, Wallace-Hanlon S, Kowalczyk OS, et al. Transcranial direct current stimulation (tDCS) combined with cognitive training in adolescent boys with ADHD: a double-blind, randomised, sham-controlled trial. *Psychol Med.* (2021). doi: 10.1017/S0033291721001859

72. Ehrhardt SE, Filmer HL, Wards Y, Mattingley JB, Dux PE. The influence of tDCS intensity on decision-making training and transfer outcomes. *J Neurophysiol.* (2021) 125. doi: 10.1152/jn.00423.2020