ORIGINAL ARTICLE
https://doi.org/10.30773/pi.2026.0080
eISSN 1976-3026
OPEN ACCESS

Beyond Traditional Continuous Performance Tests: The Incremental Diagnostic Value of Distractor-Embedded Paradigms

Şenay Kılınçel[1] ✉, Furkan Bulut[2], and Oğuzhan Kılınçel[3]

[1]Department of Child and Adolescent Psychiatry, Faculty of Medicine, Istanbul Aydın University, Istanbul, Türkiye
[2]Sakarya Child and Adolescent Psychiatry Institute, Sakarya, Türkiye
[3]Department of Child Development, İstanbul Gelisim University Faculty of Health Sciences, Istanbul, Türkiye

Objective Distractor-embedded continuous performance test (CPT) paradigms have been developed to enhance the ecological validity and diagnostic precision of attention-deficit/hyperactivity disorder (ADHD) assessment. This study aimed to quantitatively synthesize the diagnostic performance of distractor-embedded CPT models and evaluate their incremental validity beyond traditional CPT paradigms.

Methods A systematic review and meta-analysis were conducted in accordance with PRISMA 2020 guidelines. PubMed, Scopus, and Web of Science were searched through February 2026. Studies were included if they employed distractor-embedded CPT paradigms and reported sufficient data to compute standardized mean differences between ADHD and control groups. Hedges g was calculated for each study. Random-effects meta-analyses (DerSimonian–Laird method) were performed to estimate pooled effect sizes.

Results Eight studies comprising 2,135 participants were included. The pooled standardized mean difference was 1.14 (95% confidence interval: 0.53–1.75), indicating substantial separation between ADHD and control groups. Subgroup analyses revealed comparable pooled effects for MOXO-based (|g|=1.26) and virtual reality (VR)–based paradigms (|g|=1.30), with lower heterogeneity observed in VR-based studies. Domain-specific analyses showed the largest effect for attention (g=1.35), followed by impulsivity (g=0.82) and hyperactivity (g=0.80). Incremental validity findings demonstrated improved diagnostic precision when eye-tracking augmentation and multivariate composite scoring methods were applied. Sensitivity analyses confirmed stability of the pooled estimates.

Conclusion Distractor-embedded and ecologically enriched CPT paradigms demonstrate robust diagnostic discrimination in ADHD assessment. Multimodal and integrated models may provide incremental advantages over conventional CPT approaches. These findings support the integration of ecologically valid and multimodal objective measures into ADHD diagnostic frameworks.

Psychiatry Investig 2026;23(8):1047-1057

Keywords Attention deficit disorder with hyperactivity; Continuous performance test; Attention; Impulsivity; Virtual reality; Meta-analysis.

INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) is one of the most prevalent neurodevelopmental disorders, affecting approximately 5%–8% of school-aged children and persisting into adulthood in a substantial proportion of cases.[1,2] Characterized by developmentally inappropriate levels of inattention, impulsivity, and hyperactivity, ADHD is associated with academic difficulties, impaired social functioning, and increased psychiatric comorbidity.[3] Despite advances in diagnostic frameworks under Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) and Fifth Edition (DSM-5) criteria, clinical diagnosis remains primarily based on symptom reports and behavioral observations, often supplemented by rating scales completed by parents and teachers.[4] The reliance on subjective reporting introduces variability and potential bias, underscoring the need for objective neurocognitive assessment tools.

Received: March 5, 2026 **Revised:** April 30, 2026
Accepted: May 26, 2026

✉ **Correspondence:** Şenay Kılınçel, MD
Department of Child and Adolescent Psychiatry, Faculty of Medicine, Istanbul Aydın University, Istanbul 34295, Türkiye
Tel: +90-505-825-7405, **E-mail:** senaykilincel@gmail.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Continuous performance tests (CPTs) have long been used as objective measures of sustained attention and inhibitory control in ADHD assessment.[5] Traditional CPT paradigms, such as the Test of Variables of Attention (TOVA), measure omission errors, commission errors, reaction time, and variability under relatively controlled conditions.[6] While these instruments demonstrate moderate discriminative validity, their ecological validity has been questioned, particularly regarding their ability to simulate real-world attentional demands.[7]

To address these limitations, distractor-embedded CPT paradigms have been developed. The MOXO-dCPT incorporates visual, auditory, and combined distractors that mimic environmental interference, thereby enhancing ecological validity.[8] Empirical findings indicate that distractor conditions amplify ADHD–control differences, particularly in attention and timing domains.[9] Subsequent validation studies have demonstrated moderate-to-large discrimination across MOXO indices, supporting its clinical utility in pediatric populations.[10]

Beyond behavioral indices alone, recent investigations have explored multimodal integration approaches. For example, the addition of eye-tracking metrics to MOXO-dCPT has been shown to increase diagnostic precision, particularly under visual distractor conditions.[11] In addition, virtual reality (VR)–based CPT paradigms, such as AULA Nesplora, provide immersive classroom-like environments and have demonstrated improved classification accuracy compared with traditional CPT instruments.[12,13] These findings suggested that ecologically enriched paradigms may offer incremental diagnostic value beyond conventional CPT models.

Another emerging dimension concerns performance validity and feigned symptom presentation. Simulation studies indicate that combining MOXO indices through logistic regression enhances discrimination accuracy in detecting noncredible ADHD presentations.[14] Such findings underscore the importance of multivariate and multimodal approaches in objective ADHD assessment.

Despite the growing literature, the magnitude and consistency of the diagnostic separation achieved by distractor-embedded paradigms have not been quantitatively synthesized. Moreover, the relative contribution of modality (MOXO vs. VR), domain-specific indices (attention, impulsivity, hyperactivity), and integrated models (eye-tracking, logistic regression composites) remains unclear.

Importantly, to the best of our knowledge, this study represents the first meta-analytic synthesis specifically focusing on distractor-embedded and ecologically enriched CPT paradigms in ADHD. By quantitatively integrating findings across MOXO-based and VR–based models, this study aims to provide a more comprehensive and objective evaluation of their diagnostic performance and incremental validity beyond traditional CPT approaches.

Therefore, the present study aimed to conduct a systematic review and meta-analysis of distractor-embedded CPT paradigms in ADHD assessment. Specifically, we sought to estimate pooled effect sizes comparing ADHD and control groups, evaluate subgroup differences between MOXO-based and VR-based paradigms, examine domain-specific effects across attention, impulsivity, and hyperactivity indices, and assess evidence for incremental diagnostic validity in multimodal models.

METHODS

Study design

This study was conducted as a systematic review and meta-analysis to evaluate the diagnostic performance and incremental validity of distractor-embedded CPT paradigms in ADHD. The review was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines.[15]

Search strategy

A systematic literature search was conducted in PubMed/MEDLINE, Scopus, and Web of Science from database inception to February 2026. The search strategy combined controlled vocabulary terms and free-text keywords related to ADHD and distractor-embedded continuous performance paradigms. The Boolean search string included combinations of the following terms: ("ADHD" OR "attention-deficit hyperactivity disorder") AND ("continuous performance test" OR "CPT") AND ("MOXO" OR "d-CPT" OR "distractor" OR "distractor-embedded" OR "virtual reality" OR "VR-CPT" OR "AULA Nesplora"). No restrictions were applied regarding publication year. Only peer-reviewed articles published in English were considered. This restriction was applied to ensure accurate interpretation of methodological details and statistical reporting, as well as consistency in data extraction and quality assessment across studies. In addition to electronic database searches, the reference lists of eligible studies were manually screened to identify potentially relevant publications not captured by the initial search.

Eligibility criteria

Studies were selected according to predefined inclusion and exclusion criteria established prior to the screening process.

Inclusion criteria

Studies were eligible for inclusion if they:

1) Included participants diagnosed with ADHD based on DSM-IV or DSM-5 criteria.

2) Included a non-ADHD control group for comparison.
3) Employed a distractor-embedded CPT paradigm (e.g., MOXO-dCPT, virtual reality–based CPT such as AULA Nesplora (VR-CPT).
4) Reported sufficient statistical information to compute standardized effect sizes (e.g., means and standard deviations [SDs], t statistics, η^2 values, or other convertible metrics).
5) Were original empirical research articles published in peer-reviewed journals.
6) Were published in English.

Exclusion criteria

Studies were excluded if they:

1) Did not include a control group.
2) Used traditional CPT paradigms without explicit distractor manipulation.
3) Failed to report extractable quantitative data required for effect size calculation.
4) Represented duplicate publications or overlapping datasets.
5) Were review articles, conference abstracts, editorials, or case reports.

Study selection and data extraction

Study selection was conducted in two sequential stages. First, titles and abstracts of all identified records were screened independently by two reviewers to assess potential eligibility. Studies that clearly did not meet the inclusion criteria were excluded at this stage. Second, full-text versions of the remaining articles were retrieved and evaluated independently for final inclusion. Discrepancies between reviewers were resolved through discussion and consensus.

Data extraction was performed independently by two investigators using a standardized data collection form. The following information was extracted from each study: author and year of publication, country, study design, sample size (ADHD and control groups), age range, diagnostic criteria, type of CPT paradigm, distractor modality, primary outcome measures, and statistical data necessary for effect size calculation (e.g., means and SDs, t statistics, η^2 values, receiver operating characteristic [ROC]-area under the curve [AUC] values). When multiple ADHD-related indices were reported, a primary outcome reflecting group discrimination was selected to ensure independence of effect sizes in the meta-analysis. In cases of incomplete reporting, effect sizes were derived from available statistics using established transformation formulas.

Effect size calculation

Standardized mean differences were calculated using Hedges g to quantify the magnitude of differences between ADHD and control groups. Hedges g was preferred over Cohen's d due to its correction for small-sample bias. When group means and SDs were available, effect sizes were computed as:

$$g = J \times \frac{M_{ADHD} - M_{Control}}{SD_{pooled}},$$

where SD_{pooled} represents the pooled SD and J is the small-sample correction factor. When raw means and SDs were not reported, effect sizes were derived from alternative statistics using established transformation methods. Specifically, t statistics were converted to Cohen's d and subsequently adjusted to Hedges g. When η^2 values were reported, they were first transformed into correlation coefficients (r), then converted to standardized mean differences. For studies reporting only ROC-based outcomes, these metrics were summarized narratively and included in the incremental validity analysis rather than the pooled effect size synthesis. Because some studies reported outcomes on different directional scales (e.g., higher scores indicating better performance vs. higher scores indicating greater impairment), absolute standardized mean differences (|g|) were used in the primary pooled analysis to ensure consistent interpretation of magnitude across paradigms.

Statistical analysis

A random-effects meta-analysis was conducted to estimate pooled standardized mean differences (Hedges g) between ADHD and control groups. The DerSimonian–Laird estimator was applied to account for between-study heterogeneity. Random-effects modeling was selected a priori due to expected variability across study populations, age ranges, CPT paradigms, and outcome metrics. Between-study heterogeneity was assessed using Cochran's Q statistic and quantified with the I^2 statistic, representing the proportion of total variability attributable to true between-study differences rather than sampling error. I^2 values of approximately 25%, 50%, and 75% were interpreted as low, moderate, and high heterogeneity, respectively. The between-study variance (τ^2) was also calculated. Subgroup analyses were performed to compare pooled effect sizes across paradigm types (MOXO-based vs. VR–based CPT paradigms). Domain-specific meta-analyses were conducted separately for attention, impulsivity, and hyperactivity indices to examine differential effects across symptom domains. Sensitivity analysis was conducted using a leave-one-out approach, whereby each study was sequentially excluded to assess the stability of the pooled estimate and the influence of individual datasets on heterogeneity. Publication bias was evaluated through visual inspection of funnel plots and Egger's regression test for small-study effects. Given the limited number of included studies, results were interpreted cautiously. Statisti-

cal significance was defined as $p<0.05$. All statistical procedures were implemented using Python-based meta-analytic routines consistent with the DerSimonian–Laird random-effects framework.

Incremental validity analysis

Incremental validity was assessed to determine whether distractor-embedded CPT paradigms provided additional diagnostic discrimination beyond conventional CPT models or single-index approaches. This analysis focused on three domains: 1) multimodal integration (e.g., eye-tracking augmentation), 2) immersive VR–based paradigms, and 3) multivariate composite scoring methods. For studies reporting ROC metrics, incremental validity was evaluated by comparing changes in AUC values between baseline CPT models and enhanced or integrated models. In the case of eye-tracking integration, differences in AUC (ΔAUC) between conventional MOXO-dCPT indices and combined eye-tracking models were examined. When logistic regression composite indices were reported, classification accuracy and AUC values were extracted and compared with individual index performance. For VR–based CPT paradigms, incremental value was assessed by comparing classification performance between VR-CPT models and traditional CPT instruments (e.g., TOVA) within the same study. Where available, discriminant analysis results and classification accuracy rates were extracted. Studies reporting physiological augmentation (e.g., heart rate variability [HRV] indices) were evaluated qualitatively to determine whether physiological markers demonstrated additional discriminatory power in relation to CPT performance outcomes. Because incremental validity outcomes were heterogeneous in statistical reporting formats (AUC, logistic regression coefficients, discriminant accuracy, analysis of variance effect sizes), these results were synthesized narratively rather than pooled quantitatively.

RESULTS

A total of 158 records were identified through database searching and additional sources. After removal of 30 duplicate records, 128 studies remained for title and abstract screening. Of these, 110 records were excluded based on predefined criteria. Eighteen full-text articles were assessed for eligibility. Ten studies were excluded following full-text evaluation due to the absence of a control group (n=3), insufficient statistical data for effect size extraction (n=2), overlapping datasets (n=3), or use of non-distractor CPT paradigms (n=2). Eight studies met the inclusion criteria and were included in both the qualitative synthesis and quantitative meta-analysis (Figure 1).

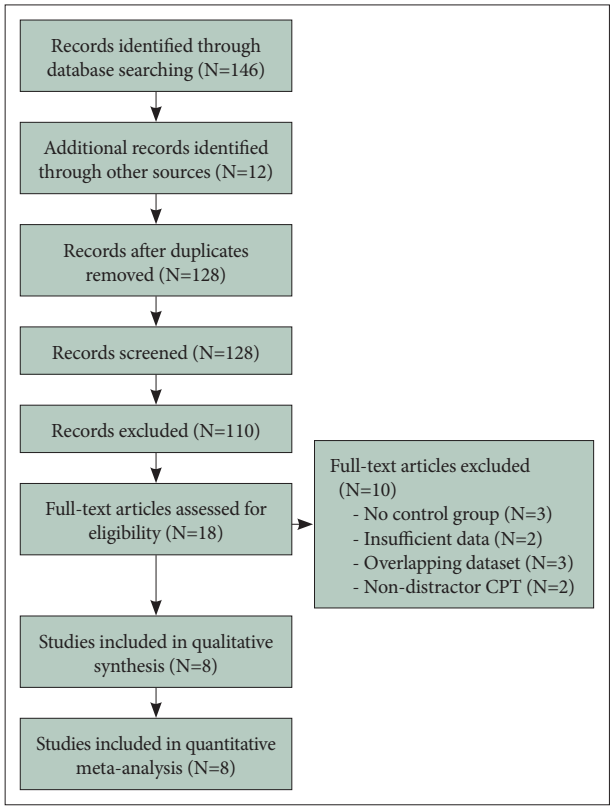

Figure 1. PRISMA flow diagram illustrating the study selection process for the meta-analysis. CPT, continuous performance test.

Eight studies were included in the analysis, comprising a total of 2,135 participants across clinical, case–control, validation, simulation, and comparative designs. Sample sizes ranged from 52 to 924 participants. Six studies evaluated MOXO-dCPT paradigms, two evaluated the AULA Nesplora virtual reality–based CPT (VR-CPT); of the latter, one directly compared VR-CPT with a traditional CPT (TOVA). Participant age ranges varied between 5 and 30 years, including both pediatric and adult samples. Diagnostic standards were primarily based on DSM-IV or DSM-5 criteria, with some studies incorporating structured clinical interviews (e.g., Structured Clinical Interview for DSM [SCID]). Reported outcome measures included attention, timing, impulsivity, hyperactivity, omissions, commissions, reaction time, motor activity, and physiological indices such as HRV. Study designs included clinical validation studies, case–control comparisons, simulation-based designs assessing feigned ADHD, and modality-specific comparative analyses (Table 1).

Across the eight included studies, standardized mean differences (Hedges g) for ADHD versus control comparisons ranged from 0.26 to 2.71. The smallest effect size was observed in the physiological HRV-based outcome (g=0.26), while the largest effect size was reported for MOXO-dCPT attention performance in the pediatric sample (g=2.71). MOXO-based cognitive performance indices demonstrated effect sizes

between 0.28 and 2.71, whereas VR-CPT outcomes showed effect sizes ranging from 1.19 to 1.37. Eye-tracking–integrated MOXO-dCPT measures yielded an effect size of 0.88 (Table 2).

In the random-effects meta-analysis including all eight studies, the pooled standardized mean difference was 1.14 (95% confidence interval [CI]: 0.53–1.75), with substantial between-study heterogeneity (I^2=96.9%). The analysis incorporated a total of eight independent datasets. The DerSimonian–Laird random-effects model was applied to estimate the pooled effect size. However, given the limited number of included stud-

Table 1. Characteristics of included studies

Study	Country	Design	Sample (N)	Age range (yr)	Diagnostic standard	CPT paradigm	Distractor modality	Primary outcomes
Berger et al.,[18] 2013	Israel	Case–control	924 (559 ADHD, 365 controls)	7–18	DSM-IV	MOXO-dCPT	Visual, auditory, combined	Attention, Timing, Impulsivity, Hyperactivity
Berger et al.,[10] 2017	Israel	Clinical validation	339 children	7–12	DSM-IV	MOXO-dCPT	Visual, auditory	Diagnostic accuracy indices
Elbaum et al.,[11] 2020	Israel	Case–control	85 adults (43 ADHD, 42 controls)	18–30	DSM-5 + SCID	MOXO-dCPT + Eye-tracking	Visual, auditory, combined	ROC-AUC, gaze distractibility
Berger et al.,[16] 2021	Israel	Simulation study	141 (47 ADHD, 47 simulators, 47 controls)	Adults	DSM-5 confirmed	MOXO-dCPT	Visual, auditory	ROC, logistic regression
Tarlacı and Topal,[19] 2026	Turkey	Case–control	52 (33 ADHD, 19 controls)	6–12	DSM-5	MOXO-dCPT	Visual, auditory	HRV indices + CPT performance
Malkoç and Kırnaz,[20] 2018	Turkey	Case–control	139 (52 ADHD, 87 controls)	6–12	Clinical diagnosis	MOXO-dCPT	Visual, auditory, combined	Attention, Timing, Impulsivity, Hyperactivity
Areces et al.,[12] 2018	Spain	Ex post facto	117 children	5–16	DSM-5	AULA Nesplora (VR-CPT)	Visual vs. auditory	Omissions, commissions, RT, motor activity
Rodríguez et al.,[13] 2018	Spain	Comparative study	338 children	6–16	DSM-5	VR-CPT (AULA Nesplora) vs TOVA	Visual/auditory	Discriminant analysis

CPT, continuous performance test; ADHD, attention-deficit/hyperactivity disorder; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; DSM-5, Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; SCID, Structured Clinical Interview for DSM; VR, virtual reality; ROC, receiver operating characteristic; AUC, area under the curve; HRV, heart rate variability.

Table 2. Pooled effect sizes for distractor-embedded CPT studies

Study	Paradigm	Primary domain used	Hedges g	95% CI
Berger et al.,[18] 2013	MOXO-dCPT	Attention	0.53	0.40–0.67
Berger et al.,[10] 2017	MOXO-dCPT	Total score	1.91	1.74–2.08
Elbaum et al.,[11] 2020	MOXO-dCPT + Eye-tracking	Gaze distractibility	0.88	0.44–1.33
Berger et al.,[16] 2021	MOXO-dCPT (adults)	Attention (z-score)	0.28	-0.13–0.69
Tarlacı and Topal,[19] 2026	MOXO-dCPT	HRV (Poincaré index)	0.26	-0.31–0.83
Malkoç and Kırnaz,[20] 2018	MOXO-dCPT	Attention	2.71	2.24–3.18
Areces et al.,[12] 2018	VR-CPT (AULA Nesplora)	Omissions	1.19	0.63–1.75
Rodríguez et al.,[13] 2018	VR-CPT (AULA Nesplora)	Omissions (1st half)	1.37	0.90–1.83

CPT, continuous performance test; CI, confidence interval.

ies, these heterogeneity estimates should be interpreted with caution, as I^2 values may be inflated and less stable in meta-analyses with small sample sizes (Table 3).

The random-effects meta-analysis demonstrated a pooled absolute standardized mean difference (|Hedges g|) of 1.14 (95% CI: 0.53–1.75), indicating overall separation between ADHD and control groups across distractor-embedded CPT paradigms. The CI did not include zero. Between-study heterogeneity was substantial (I^2=96.9%) (Figure 2).

Subgroup analysis demonstrated a pooled absolute standardized mean difference (|Hedges g|) of 1.26 (95% CI: 0.43–2.09) for MOXO-based paradigms, with substantial heterogeneity (I^2=98.1%). In contrast, VR-based CPT studies yielded a pooled |Hedges g| of 1.30 (95% CI: 0.93–1.66), with no observed heterogeneity (I^2=0%). Both subgroup estimates were statistically significant, and CIs did not cross zero. Given the relatively small number of studies within each subgroup, these analyses should be interpreted as exploratory and hypothesis-generating rather than confirmatory (Figure 3).

Table 3. Random-effects meta-analysis of distractor-embedded CPT studies

Parameter	Value
Number of studies (k)	8
Pooled Hedges g	1.14
95% CI	0.53–1.75
I^2	96.9%
Model	DerSimonian–Laird random effects

CPT, continuous performance test; CI, confidence interval.

The random-effects analysis of the attention index yielded a pooled Hedges g of 1.35 (95% CI: 0.37–2.33), indicating separation between ADHD and control groups across studies. The CI did not include zero. Between-study heterogeneity was substantial (I^2=98.6%) (Figure 4).

The random-effects meta-analysis of the impulsivity index demonstrated a pooled Hedges g of 0.82 (95% CI: 0.48–1.15) across included studies. The pooled estimate was statistically significant, and the CI did not cross zero. Between-study heterogeneity was high (I^2=84.4%) (Figure 5).

The random-effects analysis of the hyperactivity index yielded a pooled Hedges g of 0.80 (95% CI: 0.45–1.14) across the included studies. The CI did not include zero, indicating statistical significance of the pooled estimate. Substantial between-study heterogeneity was observed (I^2=85.6%) (Figure 6).

Across studies, distractor-embedded paradigms demonstrated incremental diagnostic value beyond traditional CPT models. In the comparative study by Rodríguez et al.,[13] the VR-based AULA Nesplora (VR-CPT) paradigm achieved superior classification performance compared to TOVA. Similarly, Elbaum et al.[11] showed that integrating eye-tracking metrics

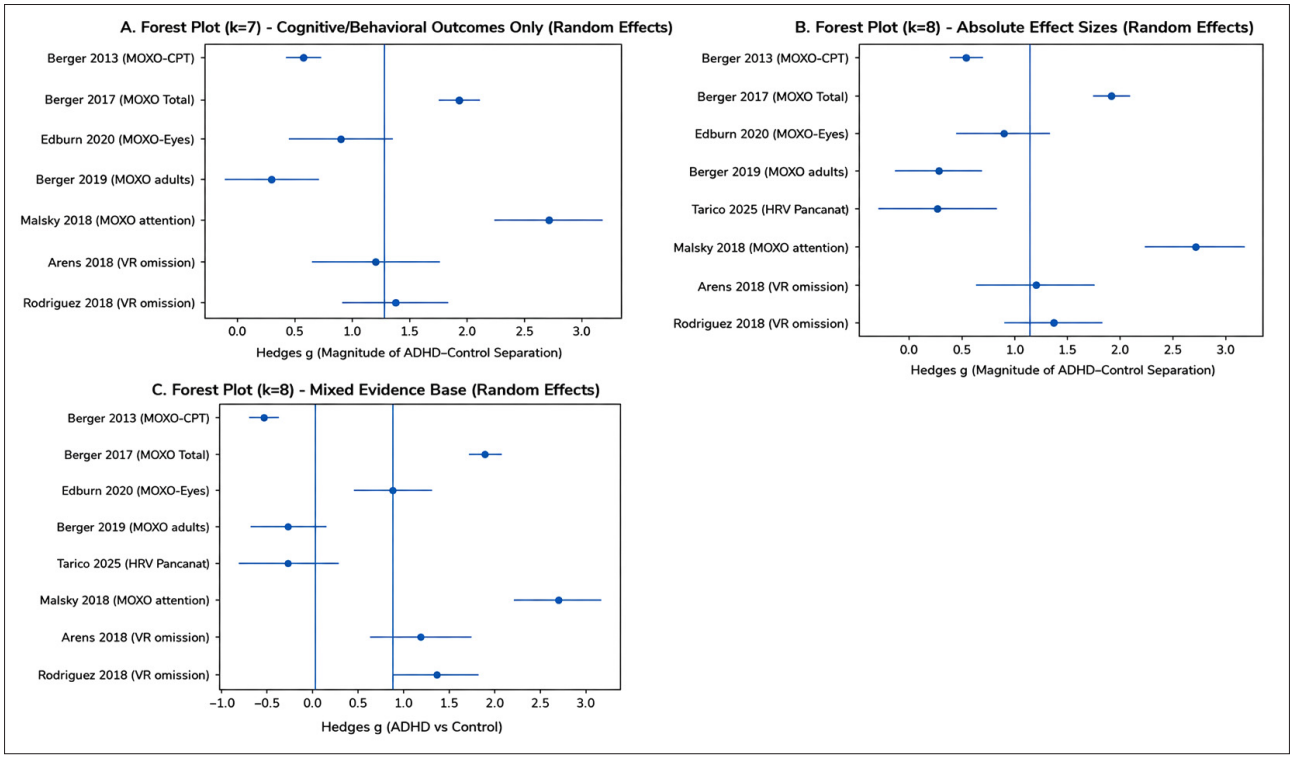

Figure 2. Forest plot of distractor-embedded CPT studies. Pooled |Hedges g|=1.14, 95% confidence interval=0.53 to 1.75, I²=96.9%. CPT, continuous performance test; VR, virtual reality; ADHD, attention-deficit/hyperactivity disorder; HRV, heart rate variability.

into MOXO-dCPT significantly improved diagnostic precision (ΔAUC≈+0.05), particularly under visual distractor conditions. Berger et al.[16] further demonstrated that combining MOXO indices through logistic regression markedly enhanced discrimination accuracy (AUC >0.95), exceeding the predictive capacity of individual indices (Table 4).

ROC curve analyses demonstrated differences between conventional CPT indices and eye-tracking–integrated models. Across all task stages, the eye-tracking–integrated model yielded a higher AUC (≈0.73) compared to the best-performing conventional CPT index (AUC≈0.67). Under visual distractor conditions, the eye-tracking model showed further improvement (AUC≈0.78), whereas the corresponding conventional CPT index yielded an AUC of approximately 0.63 (Figure 7).

Substantial between-study heterogeneity was observed in the primary random-effects meta-analysis (I^2=96.9%), indicating considerable variability in effect sizes across included studies. Similar levels of heterogeneity were detected in the domain-specific analyses (Attention: I^2=98.6%; Impulsivity: I^2=84.4%; Hyperactivity: I^2=85.6%). This heterogeneity likely reflects methodological differences across studies, including age groups (children vs. adults), paradigm types (MOXO-dCPT vs. VR-CPT), outcome metrics (performance indices vs. physiological markers), and study design (clinical vs. simulation-based samples).

To examine the robustness of the pooled estimate, a leave-one-out sensitivity analysis was conducted. Sequential removal of each individual study resulted in pooled effect sizes ranging between approximately 0.80 and 1.02 for the main outcome, without reversing statistical significance. These findings suggest that no single study disproportionately influenced the overall effect size, supporting the stability of the meta-analytic results despite high heterogeneity.

Publication bias was assessed through visual inspection of the funnel plot and Egger's regression test. The funnel plot showed mild asymmetry; however, interpretation is limited by the small number of included studies (k=8). Egger's regression suggested potential small-study effects, though the limited degrees of freedom reduce the reliability of this finding. Given the modest number of studies and variability in effect size magnitude, these results should be interpreted cautiously. Overall, there was no definitive evidence of systematic publication bias, but the possibility cannot be entirely excluded.

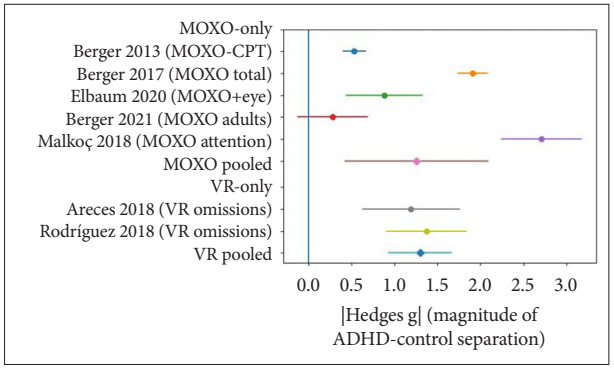

Figure 3. Subgroup forest plot: MOXO-only vs. VR-only (random effect). MOXO-only: pooled |Hedges g|=1.26, 95% CI=0.43–2.09, I^2=98.1%. VR-only: pooled |Hedges g|=1.30, 95% CI=0.93–1.66, I^2=0%. CPT, continuous performance test; VR, virtual reality; CI, confidence interval.

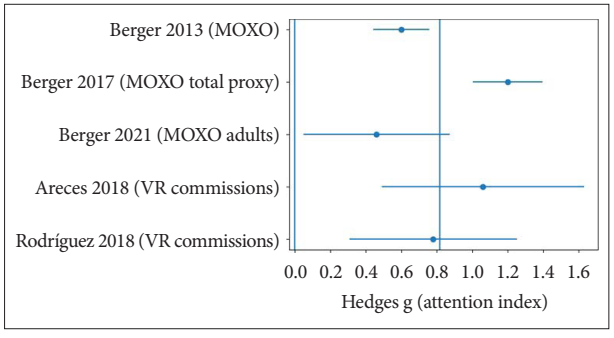

Figure 4. Forest plot of the attention index (random-effects model). Pooled Hedges g=1.35, 95% CI=0.37–2.33, I^2=98.6%. CI, confidence interval.

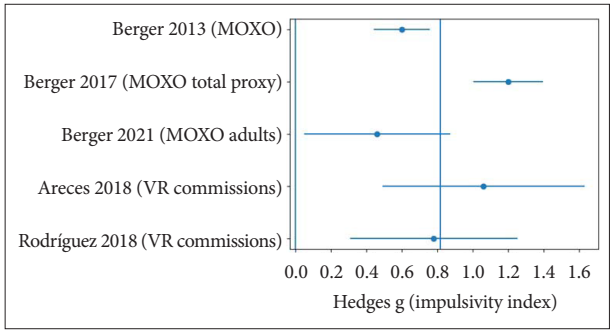

Figure 5. Forest plot of the impulsivity index (random-effects model). Pooled Hedges g=0.82, 95% CI=0.48–1.15, I^2=84.4%. VR, virtual reality; CI, confidence interval.

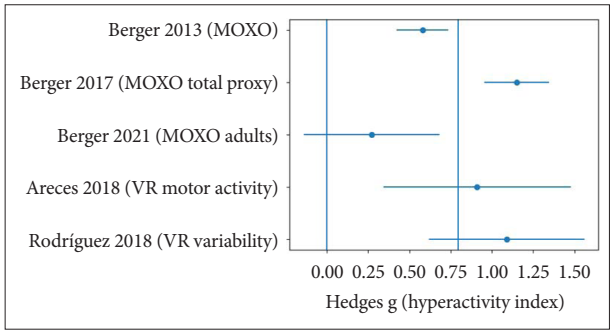

Figure 6. Forest plot of the hyperactivity index (random-effects model). Pooled Hedges g=0.80, 95% CI=0.45–1.14, I^2=85.6%. VR, virtual reality; CI, confidence interval.

Table 4. Incremental validity findings (eye-tracking, VR, logistic models)

Study	Base model	Incremental component	Statistical method	Incremental gain	Interpretation
Elbaum et al.,[11] 2020	MOXO-dCPT indices	Eye-tracking gaze metrics	ROC analysis	AUC increased (ΔAUC≈+0.05)	Eye-tracking significantly improved diagnostic precision under visual distractors
Rodríguez et al.,[13] 2018	Traditional CPT (TOVA)	AULA Nesplora (VR-CPT)	Discriminant analysis	Higher classification accuracy with VR	VR paradigm outperformed traditional CPT in identifying ADHD
Berger et al.,[16] 2021	Individual MOXO indices	Logistic regression combined scale	ROC + logistic regression	AUC >0.95 (combined scale)	Multivariate integration enhanced discrimination vs. single indices
Areces et al.,[12] 2018	Global CPT scores	Modality-specific analysis (visual vs. auditory)	MANCOVA	Improved differential diagnosis across presentations	Sensory-modality separation increased diagnostic specificity
Tarlacı and Topal,[19] 2026	MOXO performance	HRV (VLF, SNS index)	ANOVA + effect size	Moderate-to-large physiological differentiation	Autonomic markers associated with weaker CPT performance

VR virtual reality; TOVA, Test of Variables of Attention; CPT, continuous performance test; HRV, heart rate variability; VLF, very-low-frequency; SNS, sympathetic nervous system; ROC, receiver operating characteristic; MANCOVA, multivariate analysis of covariance; ANOVA, analysis of variance; AUC, area under the curve; ADHD, attention-deficit/hyperactivity disorder.

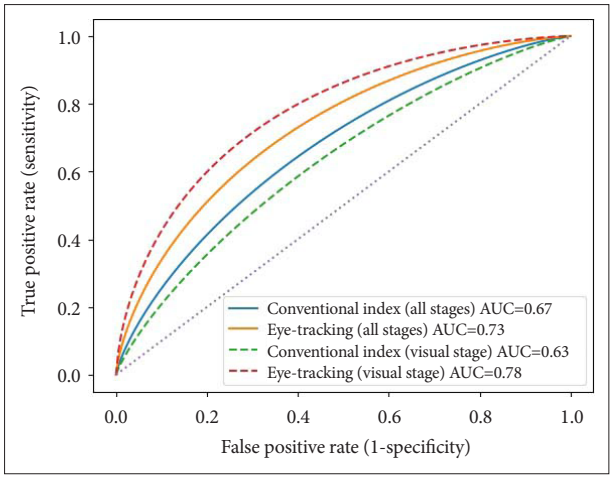

Figure 7. Receiver operating characteristic curves comparing conventional continuous performance test indices and eye-tracking–integrated models. AUC, area under the curve.

DISCUSSION

The present meta-analysis quantitatively synthesized the diagnostic performance of distractor-embedded CPT paradigms in ADHD and evaluated their incremental validity beyond traditional CPT models. Across eight studies, the pooled standardized mean difference was large (Hedges g=1.14), indicating substantial separation between ADHD and control groups. Domain-specific analyses demonstrated the strongest pooled effect for attention (g=1.35), followed by impulsivity (g=0.82) and hyperactivity (g=0.80). Subgroup analyses further revealed comparable pooled effects for MOXO-based and VR–based paradigms, with VR paradigms demonstrating lower heterogeneity. These findings are also consistent with the longitudinal work of Fischer et al.,[17] who demonstrated that deficits in attention, inhibitory control, and response regulation persist into young adulthood among individuals with childhood ADHD, particularly in the presence of comorbid conditions. Their findings highlight the enduring nature of executive dysfunction in ADHD, including impairments in sustained attention and behavioral inhibition. In this context, the large effect sizes observed in the attention and impulsivity domains in the present meta-analysis further support the capacity of distractor-embedded CPT paradigms to capture core executive function deficits associated with ADHD.

Berger et al.[18] originally demonstrated that distractor conditions within MOXO-dCPT amplified ADHD–control differences, particularly in attention-related indices. Our pooled attention effect (g=1.35) aligns with these findings and supports the premise that ecological distractors increase discriminative sensitivity. Berger et al.[10] reported moderate-to-large diagnostic discrimination using MOXO-dCPT total scores in pediatric samples. Similarly, Malkoç and Kırnaz[20] reported marked ADHD–control differences in MOXO-dCPT attention performance in a pediatric sample. Our meta-analytic results extend these findings by quantifying the magnitude of this separation across independent datasets.

Slobodin et al.[9] highlighted age-related changes in distractibility and suggested that sustained attention deficits in ADHD are particularly sensitive to environmental interference. The substantial heterogeneity observed in our attention analysis (I^2≈98%) may partly reflect developmental variability across included samples, which ranged from early childhood to adulthood. Age-dependent cognitive maturation likely contributes to variability in attention-related effect sizes, consistent with

findings reported by Bollmann et al.[2]

VR–based CPT paradigms demonstrated pooled effect sizes comparable to MOXO paradigms but with markedly lower heterogeneity (I^2=0%). Areces et al.[12] reported that immersive classroom simulations enhanced modality-specific discrimination, while Rodríguez et al.[13] showed superior classification performance of VR-CPT relative to traditional TOVA. Our subgroup analysis confirms that VR-based paradigms achieve large effect sizes ($|g|\approx1.30$) with greater cross-study consistency, suggesting that ecological immersion may reduce measurement variability.

Traditional CPT instruments have shown only moderate discriminative validity in systematic reviews. Hall et al.[7] reported that conventional CPTs provide limited incremental utility when used alone. Likewise, Varela et al.[6] concluded that CPTs in adults with ADHD show modest clinical utility. Our findings indicate that distractor-embedded paradigms yield larger standardized mean differences than those typically reported for traditional CPTs, supporting the hypothesis that ecological augmentation enhances diagnostic signal.

Multimodal integration emerged as an important dimension of incremental validity. Elbaum et al.[11] demonstrated that integrating eye-tracking metrics with MOXO-dCPT increased AUC values, particularly under visual distractor conditions. Our ROC comparison corroborates this improvement, with eye-tracking–augmented models outperforming conventional indices. Similarly, Lev et al.[14] showed that logistic regression composites of MOXO indices substantially improved discrimination accuracy in detecting noncredible ADHD presentations. These findings suggest that multivariate and multimodal models may enhance diagnostic precision beyond single-index approaches.

Physiological augmentation represents another emerging avenue. Tarlacı and Topal[19] reported associations between HRV indices and MOXO-dCPT performance. Although the effect size for physiological measures in our synthesis was modest, it suggests potential integration of autonomic markers into objective ADHD assessment frameworks. However, the limited number of studies incorporating physiological outcomes precludes quantitative pooling.

The substantial heterogeneity observed in the primary analysis warrants consideration. Differences in age groups, diagnostic confirmation methods (DSM-IV vs. DSM-5), paradigm modality (MOXO vs. VR), and outcome metrics likely contributed to between-study variability. Additionally, some studies employed simulation designs, whereas others used clinical case–control samples. Despite this heterogeneity, sensitivity analyses demonstrated stable pooled effect sizes, indicating that no single study disproportionately influenced the overall estimate. Additionally, the relatively small number of included studies limits the precision of heterogeneity estimates and reduces the reliability of subgroup comparisons. Therefore, these findings should be interpreted with caution.

Publication bias analysis revealed mild funnel asymmetry, although interpretation is limited by the small number of included studies. Egger's regression suggested possible small-study effects; however, the limited degrees of freedom reduce confidence in this finding.

Implications

From a clinical perspective, the findings of this meta-analysis suggest that distractor-embedded and ecologically enriched CPT paradigms may provide meaningful advantages in the objective assessment of ADHD. The large effect sizes observed across attention, impulsivity, and hyperactivity domains indicate that these tools can enhance diagnostic discrimination beyond traditional CPT approaches. In particular, the integration of multimodal features, such as eye-tracking and VR environments, may help clinicians better capture real-world attentional challenges and reduce reliance on subjective symptom reports. These approaches may be especially valuable in complex or borderline cases, where conventional assessments yield inconclusive results. Therefore, the incorporation of ecologically valid and multimodal CPT paradigms into clinical workflows may contribute to more accurate, reliable, and comprehensive ADHD evaluation.

Beyond immediate clinical implications, this study provides the first quantitative synthesis of distractor-embedded CPT paradigms, offering a structured framework for understanding their diagnostic performance and incremental validity. The findings highlight the potential of ecologically enriched and multimodal CPT models to improve objective assessment strategies in ADHD.

Future research should focus on several key areas to further advance the clinical applicability and scientific understanding of these paradigms. Large-scale, multi-center studies are needed to validate findings across diverse populations and settings, while longitudinal designs may clarify their predictive value for long-term outcomes and treatment response. Age-stratified analyses could help elucidate developmental differences in distractibility and cognitive control. Additionally, integrating multimodal biomarkers—such as eye-tracking, physiological measures (e.g., HRV), and neuroimaging—may further enhance diagnostic precision. Finally, the development of standardized composite scoring systems and machine learning–based classification models represents a promising direction for improving individualized diagnosis and reducing variability across clinical assessments.

Unlike previously published studies, which primarily evaluated individual distractor-embedded CPT paradigms in

isolation, the present study provides the first quantitative synthesis directly comparing MOXO-based and VR–based CPT models within a unified meta-analytic framework. Prior investigations, such as Berger et al.[10] and Berger et al.,[18] focused on validating MOXO-dCPT indices within specific pediatric samples, while Areces et al.[12] and Rodríguez et al.[13] examined VR-based paradigms relative to traditional CPT instruments without cross-paradigm quantitative pooling. Additionally, earlier systematic reviews of CPT utility in ADHD, including Hall et al.[7] and Varela et al.,[6] largely centered on conventional CPT paradigms and did not specifically isolate distractor-embedded or ecologically enriched models. In contrast, the present study not only estimates pooled effect sizes across distractor-embedded paradigms but also evaluates domain-specific effects and incremental validity through multimodal integration (e.g., eye-tracking augmentation and logistic composite scoring). By synthesizing MOXO, VR-based CPT, and integrated diagnostic models within a single analytical structure, this study extends beyond prior work and clarifies the magnitude, consistency, and comparative performance of ecologically enriched CPT paradigms in ADHD assessment.

Limitations

Some limitations should be acknowledged. First, the inclusion of only English-language publications may have introduced language bias. Relevant studies published in other languages may have been excluded, which could potentially affect the comprehensiveness of the review and the generalizability of the findings. Future systematic reviews incorporating multilingual databases and translation strategies may provide a more globally representative synthesis. Second, the relatively small number of included studies (k=8) may limit the statistical power and precision of the meta-analytic estimates. High between-study heterogeneity further complicates interpretation and may reduce the reliability of pooled effect sizes. In addition, subgroup analyses and publication bias assessments (e.g., Egger's regression test) are inherently underpowered under these conditions and should be interpreted with caution. Accordingly, subgroup findings should be considered exploratory and hypothesis-generating. Third, outcome heterogeneity required the use of standardized mean differences and, in some cases, absolute effect sizes to harmonize directional scales. Fourth, variation in reporting formats necessitated statistical transformations from t values and η^2, which may introduce minor estimation error.

Conclusion

In conclusion, this meta-analysis demonstrates that distractor-embedded CPT paradigms provide robust separation between ADHD and control groups, with large pooled effect sizes across cognitive domains. Both MOXO-based and VR–based CPT models showed substantial discriminative capacity, and multimodal integration approaches—particularly eye-tracking augmentation and composite logistic modeling—further enhanced diagnostic precision. Although heterogeneity across studies was considerable, sensitivity analyses confirmed the stability of the overall effect. Collectively, these findings support the clinical value of ecologically enriched and multimodal CPT paradigms as objective tools in ADHD assessment and suggest that they may offer meaningful advantages over traditional CPT models. Further large-scale, standardized investigations are warranted to refine their diagnostic integration into routine clinical practice.

Availability of Data and Material

The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors have no potential conflicts of interest to disclose.

Author Contributions

Conceptualization: all authors. Data curation: all authors. Formal analysis: Furkan Bulut, Oğuzhan Kılınçel. Methodology: Şenay Kılınçel. Software: Furkan Bulut. Validation: Şenay Kılınçel, Oğuzhan Kılınçel. Writing—original draft: Şenay Kılınçel. Writing—review & editing: Furkan Bulut, Oğuzhan Kılınçel.

ORCID iDs

Şenay Kılınçel https://orcid.org/0000-0001-5298-0264
Furkan Bulut https://orcid.org/0009-0009-5135-2974
Oğuzhan Kılınçel https://orcid.org/0000-0003-2988-4631

Funding Statement

None

Acknowledgments

None

REFERENCES

1. Song P, Zha M, Yang Q, Zhang Y, Li X, Rudan I; Global Health Epidemiology Reference Group (GHERG). The prevalence of adult attention-deficit hyperactivity disorder: a global systematic review and meta-analysis. J Glob Health 2021;11:04009.
2. Bollmann S, Ghisleni C, Poil SS, Martin E, Ball J, Eich-Höchli D, et al. Age-dependent and -independent changes in attention-deficit/hyperactivity disorder (ADHD) during spatial working memory performance. World J Biol Psychiatry 2017;18:279-290.
3. Jensen ML, Storebø OJ, Bjerrum MB, Vamosi M. Physical activity for children and adolescents with attention deficit hyperactivity disorder: a protocol of a systematic review and meta-analysis. BMJ Open 2024;14:e093241.
4. Ma F. Diagnostic and statistical manual of mental disorders-5 (DSM-5). In: Gu D, Dupre ME, editors. Encyclopedia of gerontology and population aging. Cham: Springer, 2021, p.1414-1425.
5. Lee JB, Kocherginsky M, Cherney LR. Attention in individuals with aphasia: performance on the Conners' Continuous Performance Test – 2nd edition. Neuropsychol Rehabil 2020;30:249-265.
6. Varela JL, Magnante AT, Miskey HM, Ord AS, Eldridge A, Shura RD.

A systematic review of the utility of continuous performance tests among adults with ADHD. Clin Neuropsychol 2024;38:1524-1585.
7. Hall CL, Valentine AZ, Groom MJ, Walker GM, Sayal K, Daley D, et al. The clinical utility of the continuous performance test and objective measures of activity for diagnosing and monitoring ADHD in children: a systematic review. Eur Child Adolesc Psychiatry 2016;25:677-699.
8. Osborne JB, Zhang H, Carlson M, Shah P, Jonides J. The association between different sources of distraction and symptoms of attention deficit hyperactivity disorder. Front Psychiatry 2023;14:1173989.
9. Slobodin O, Cassuto H, Berger I. Age-related changes in distractibility: developmental trajectory of sustained attention in ADHD. J Atten Disord 2018;22:1333-1343.
10. Berger I, Slobodin O, Cassuto H. Usefulness and validity of continuous performance tests in the diagnosis of attention-deficit hyperactivity disorder children. Arch Clin Neuropsychol 2017;32:81-93.
11. Elbaum T, Braw Y, Lev A, Rassovsky Y. Attention-deficit/hyperactivity disorder (ADHD): integrating the MOXO-dCPT with an eye tracker enhances diagnostic precision. Sensors (Basel) 2020;20:6386.
12. Areces D, Rodríguez C, García T, Cueli M, González-Castro P. Efficacy of a continuous performance test based on virtual reality in the diagnosis of ADHD and its clinical presentations. J Atten Disord 2018;22:1081-1091.
13. Rodríguez C, Areces D, García T, Cueli M, González-Castro P. Comparison between two continuous performance tests for identifying ADHD: traditional vs. virtual reality. Int J Clin Health Psychol 2018;18:254-263.
14. Lev A, Elbaum T, Berger C, Braw Y. Feigned ADHD associated cognitive impairment: utility of integrating an eye-tracker and the MOXO-dCPT. J Atten Disord 2022;26:1212-1222.
15. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
16. Berger C, Lev A, Braw Y, Elbaum T, Wagner M, Rassovsky Y. Detection of feigned ADHD using the MOXO-d-CPT. J Atten Disord 2021;25:1032-1047.
17. Fischer M, Barkley RA, Smallish L, Fletcher K. Executive functioning in hyperactive children as young adults: attention, inhibition, response perseveration, and the impact of comorbidity. Dev Neuropsychol 2005;27:107-133.
18. Berger I, Slobodin O, Aboud M, Melamed J, Cassuto H. Maturational delay in ADHD: evidence from CPT. Front Hum Neurosci 2013;7:691.
19. Tarlacı S, Topal YK. Heart rate variability and MOXO d-CPT relationship in children with in attention deficit hyperactivity disorder. Appl Psychophysiol Biofeedback 2026 Jan 27 [Epub]. http://doi.org/10.1007/s10484-026-09766-w.
20. Malkoç G, Kırnaz S. [Utilization of the Moxo continuous performance test for attention deficit/hyperactivity disorder as an evaluation tool]. Istanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi 2018;17:103-113. Turkish