The effects of combined action observation and motor imagery on corticospinal excitability and movement outcomes: A meta-analysis

Samantha Chye [a*], Ashika Chembila Valappil [a*], David J. Wright [b], Cornelia Frank [c], David A. Shearer [d], Christopher J. Tyler [a], Ceri E. Diss [a], Omar S. Mian [a], Neale A. Tillin [a] & Adam M. Bruton [a**]

[a] Centre for Integrated Research in Life and Health Sciences, School of Life and Health Sciences, University of Roehampton, London, UK, SW15 4JD. E-mail: chyes1@roehampton.ac.uk, chembila@roehampton.ac.uk, chris.tyler@roehampton.ac.uk, c.diss@roehampton.ac.uk, omar.mian@roehampton.ac.uk, neale.tillin@roehampton.ac.uk, adam.bruton@roehampton.ac.uk.

[b] Research Centre for Health, Psychology and Communities, Department of Psychology, Manchester Metropolitan University, Manchester, UK, M15 6GX. E-mail: d.j.wright@mmu.ac.uk.

[c] Sports and Movement Research Group, Department of Sports and Movement Science, School of Educational and Cultural Studies, Osnabrück University, Osnabrück, Germany. Email: cornelia.frank@uni-osnabrueck.de.

[d] Faculty of Life Science and Education, University of South Wales, UK, CF37 1DL. Email: david.shearer@southwales.ac.uk.

Author Note: *Samantha Chye and Ashika Chembila Valappil are shared first authors for this manuscript. **Adam M. Bruton is the corresponding author for this manuscript: adam.bruton@roehampton.ac.uk, Room 1051 Parkstead House, Whitelands College, School of Life and Health Sciences, University of Roehampton.

Abstract

Motor simulation interventions involving motor imagery (MI) and action observation (AO) have received considerable interest in the behavioral sciences. A growing body of research has focused on using AO and MI simultaneously, termed 'combined action observation and motor imagery' (AOMI). The current meta-analysis collated and synthesized existing motor evoked potential (MEP) amplitude data from transcranial magnetic stimulation studies and movement outcome data from behavioral studies to quantify changes in corticospinal excitability and motor skill performance for AOMI compared to AO, MI and control conditions. AOMI had a medium positive effect on MEP amplitudes compared to all three conditions. However, for movement outcome data, AOMI had a small to medium positive effect compared to control, but no effect compared to AO and MI conditions. The timing and intensity of transcranial magnetic stimulation moderated facilitation of corticospinal excitability during AOMI, whereas self-reported kinesthetic imagery ability moderated the effects of AOMI on movement outcomes. The results of this meta-analysis are discussed in relation to existing literature on motor simulation and skill acquisition. The limitations within the current literature on AOMI are considered, before providing viable directions for future research on this topic.

Keywords: AOMI, motor evoked potentials, dual action simulation, motor execution, motor imagery during action observation, transcranial magnetic stimulation

1. Introduction

According to motor simulation theory (Jeannerod, 1994, 2001, 2006), it is possible to cognitively rehearse an action both overtly and covertly through action observation (AO) and motor imagery (MI), with this cognitive simulation activating motor regions of the brain in a similar manner to physical execution of the action. AO is a bottom-up process that involves the deliberate and structured observation of human movement (Neuman & Gray, 2013), whereas MI is a top-down process that involves the internal generation of the visual and kinesthetic elements of movement (Macintyre et al., 2013). Literature has reported positive behavioral outcomes for AO- and MI-based practice in sport (e.g., Guillot & Collet, 2008; Ste-Marie et al., 2012) and neurorehabilitation settings (e.g., Buccino, 2014; De Vries & Mulder, 2007). There is preliminary evidence indicating that plastic changes in the primary motor system may underpin these behavioral improvements (e.g., Yoxon and Welsh, 2020). A recent large-scale meta-analysis of functional magnetic resonance imaging (fMRI) data reported a shared network including premotor, rostral parietal, and somatosensory areas of the brain active during AO, MI, and movement execution (Hardwick et al., 2018). Notably, there were differences in the neural regions activated during AO and MI, and the brain activity for these overlapped differently with brain activity for physically performed actions. It is possible that using AO and MI simultaneously, typically labelled 'combined action observation and motor imagery' (AOMI), may hold greater neural overlap with physical execution.

AOMI refers to a person watching a video or live demonstration of a movement while simultaneously generating, maintaining, and transforming a time-synchronized kinesthetic representation of the same action (Eaves, Riach et al., 2016; Vogt et al., 2013). AOMI has received growing research interest over the last decade and two hypotheses have been proposed to explain why AOMI may be more effective as a motor skill intervention than independent AO or MI interventions. First, Eaves and colleagues (2012, 2014, 2016) suggested the dual action simulation hypothesis (DASH), which proposes that a person will generate separate motor representations for the observed and imagined actions and maintain these as two parallel sensorimotor streams when

they engage in AOMI. If a person is simultaneously observing and imagining the same action, these two motor representations are likely to merge as one sensorimotor stream, producing more widespread activity in the premotor cortex compared to AO or MI alone. This is likely due to AOMI increasing activity in shared brain areas for AO and MI, as well as increasing activity in areas solely recruited during AO (e.g., inferior frontal gyrus, ventral premotor area) and MI (e.g., angular gyrus, dorsal premotor area) of an action (Filimon et al., 2015; Hardwick et al., 2018). Second, Meers et al. (2020) introduced the visual guidance hypothesis (VGH) as an alternative account of how AOMI may influence action. They suggest that MI is prioritized during AOMI, and that the AO component might merely serve as an external visual guide that facilitates more vivid MI generation. In contrast to DASH, this would mean that AO does not activate a separate motor representation during AOMI, but rather strengthens the motor representation resulting from MI. Irrespective of the stance taken, both the DASH and VGH suggest that AOMI has the capacity to influence motor skill execution above and beyond AO or MI in isolation through increased activity in motor regions of the brain.

Current neuroscientific evidence using a range of modalities supports this notion, as cortico-motor activity is increased during AOMI of an action compared to independent AO or MI of that same action (Eaves, Riach et al., 2016). Studies using fMRI report distinct neural signatures for AO, MI, and AOMI whereby the blood-oxygen-level-dependent (BOLD) signal is increased and more widespread in the brain regions involved in movement execution when an individual engages in AOMI (e.g., Nedelko et al., 2012; Taube et al., 2015; Villiger et al., 2013). For example, Taube et al. (2015) found greater activation in the supplementary motor area, basal ganglia and cerebellum during AOMI compared to AO, and greater bilateral activity in the cerebellum and greater activation in the precuneus compared to MI. Studies using electroencephalography (EEG) report that AOMI leads to significantly larger event-related desynchronization in the mu/alpha and beta frequency bands, indicative of increased activity over the primary sensorimotor areas of the brain compared to both AO and MI alone (Berends et al., 2013; Eaves, Behmer et al., 2016). These fMRI and EEG findings have important implications for applied practice, where the use of AOMI may prove

beneficial in reinforcing motor (re)learning. The increased neural activity during AOMI has the potential to support repetitive Hebbian modulation of intracortical and subcortical excitatory mechanisms through synaptic plasticity, in a similar manner to physical practice, and thus may be an effective method for behavior change (Holmes & Calmels, 2008).

From a neurophysiological perspective, this meta-analysis will focus on studies adopting single-pulse transcranial magnetic stimulation (TMS) during AOMI, as this is the most prevalent neuroscientific modality adopted in the literature to date (study count: n = 19 TMS, n = 14 fMRI, n = 10 EEG). When applied to a muscle representation of the primary motor cortex, TMS produces a twitch response in the corresponding muscles called motor evoked potentials (MEPs). MEPs are measured using surface electromyography (EMG) and provide a marker of corticospinal excitability during the time of stimulation (Rothwell, 1997). Studies using TMS during AOMI have explored changes in MEP amplitudes across a range of movements, including finger movements (Bruton et al., 2020), basketball free throws (Wright, Wood et al., 2018a), walking (Kaneko et al., 2018), and balance movements (Mouthon et al., 2016). Current literature predominantly shows increased corticospinal excitability during AOMI compared to baseline conditions (e.g., Bruton et al., 2020; Wright et al., 2014, 2016). However, studies comparing AOMI against AO or MI have reported increased (e.g., Mouthon et al., 2016; Wright et al., 2016), as well as no differences (e.g., Castro et al., 2021; Mouthon et al., 2015) in corticospinal excitability during AOMI. Given the prevalence of TMS studies exploring the neurophysiological mechanisms underpinning AOMI engagement, it is now possible to synthesize the available MEP amplitude data to quantify the effects of AOMI on corticospinal excitability compared to AO, MI, and control conditions.

AOMI investigations have explored a range of movement outcomes and types, such as movement time for ball rotations (Kawasaki et al., 2018), upper-limb kinematics for dart throwing (Romano-Smith et al., 2019), force production in Nordic hamstring curls (Scott et al., 2018), and radial error from the hole in golf-putting (Marshall et al., 2016). This research has been conducted in sport (e.g., Romano-Smith et al., 2019) and rehabilitation (e.g., Scott et al., 2018) contexts, with

neurotypical (e.g., Di Rienzo et al., 2019) and neurodivergent (e.g., Marshall et al., 2020) populations. The existing literature has almost exclusively demonstrated that movement outcomes are improved for different motor skills after AOMI interventions when compared to control conditions (e.g., Marshall et al., 2019; Romano-Smith et al., 2018; Shimada et al., 2019). However, comparisons with AO and MI interventions are equivocal, with some studies showing greater improvements in movement outcomes for AOMI compared to AO (e.g. Bek et al., 2019) and MI interventions (e.g. Scott et al., 2018), and other studies showing no such effects (e.g., Romano-Smith et al., 2019) or greater improvements for MI-related interventions (e.g., Marshall et al., 2016). Therefore, it is unclear if AOMI should be recommended as the optimal simulation approach when attempting to improve motor skill performance, highlighting the need to synthesize available movement outcome data for AOMI interventions.

Early reviews on AOMI (Eaves, Riach et al., 2016; Vogt et al., 2013) summarized the behavioral neuroscience literature contrasting AO and MI and drew on early AOMI research to support its use as a motor skill intervention. Since then, research has explored changes in MEP amplitudes and movement outcomes associated with engagement in AOMI. This has led to population-specific reviews outlining how AOMI can be used to address sensorimotor deficits for individuals with Parkinson's disease (Caligiore et al., 2017), developmental coordination disorder (Scott et al., 2021), or during post-stroke rehabilitation (Emerson et al., 2018). Systematic reviews and meta-analyses have synthesized the respective effects of MI and AO on corticospinal excitability (e.g., Grosprêtre et al., 2016; Naish et al., 2014) and motor skill performance (e.g., Ashford et al., 2006; Simonsmeier et al., 2020, Toth et al., 2020) when used in isolation, however, to our knowledge, no such meta-analyses exist for AOMI interventions.

The purpose of the current meta-analysis was to methodically collate and synthesize the available MEP amplitude data as an indicator of corticospinal excitability during AOMI engagement, and movement outcome data to assess changes in motor skill performance that result from AOMI interventions. Our primary aim was to establish the effectiveness of AOMI by comparing its effects

on MEP amplitudes and movement outcomes against those for AO, MI, and control conditions. Based on previous literature (see Eaves, Riach et al., 2016), it was hypothesized that AOMI would have a small positive effect when compared to independent AO or MI, and a medium positive effect when compared to control conditions for both outcome variables. Our secondary aim was to explore several methodological parameters hypothesized to have a moderating effect on the impact of AOMI interventions on MEP amplitudes or movement outcomes. Considering the influence of these methodological aspects directly addressed questions raised in early reviews on AOMI (Eaves, Riach et al., 2016; Vogt et al., 2013) and may permit optimal delivery of AOMI interventions in the future.

2. Method

2.1. Study Protocol

The study procedures were performed according to the methodological guidelines highlighted by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Page et al., 2020). The meta-analysis justification and proposed procedures were documented in a pre-registration protocol document (https://osf.io/9yebv) that is stored alongside additional supplementary files (https://osf.io/9yebv) on the Open Science Framework. A PRISMA flow chart (Figure 1) outlines the study search and selection process for the current meta-analysis.

[INSERT FIGURE 1 HERE]

2.1.1. Literature Search and Selection

The literature search was performed using three online databases: PubMed, Web of Science, and PsycINFO. The original search was run in August 2020 and the final search was run in June 2021. Consistent search terms were decided upon and adapted for each database based on the search string requirements (see supplementary files for full search string information). Web of Science was used as an initial search point, followed by PubMed and PsycINFO to cover the biomedical and psychological literature, respectively. The search was limited to studies with human populations that had been published in a peer-reviewed journal in English language, without any limitations to publication years. The literature search with the respective keyword combinations and restrictions

provided 396 hits in total, after 244 duplicates were removed. An exploratory search of review articles and prior knowledge of research led to 8 additional papers being added, which took the final number of papers being screened to 404.

2.1.2. Inclusion of Studies

The decision to include a study was based on five criteria. (1) The study had to either test the effects of AOMI on corticospinal excitability by recording MEP amplitudes during single pulse TMS or test the effects of AOMI on movement outcomes by recording parameters related to motor task performance. Participants were required to engage in an AOMI task as their intervention or experimental method, and AOMI had to be delivered in a synchronous manner (i.e., engagement in both forms of simulation at the same time). Studies with independent AO or MI conditions were included, but only if there was a comparison with an AOMI condition. Studies that included AOMI as an adjunct to physical practice were also included, but the inclusion of physical practice was not mandatory, and these studies needed to incorporate physical practice in the control condition to permit assessment of AOMI effects. (2) The study had to adopt a controlled experimental design where the unconfounded effects of the AOMI intervention could be compared within- or between-groups. As such, studies using a correlational or survey design, qualitative research, or case study/single-subject designs were excluded. Studies that did not report a control group or baseline condition were also excluded, unless they contained data from AO or MI conditions that could be compared to the AOMI condition. (3) The study had to include an MEP amplitude or movement outcome measure with a specific experimental group, or across all study groups. Studies that did not have distinguishable MEP amplitude or movement outcome measures or used neurophysiological modalities other than/combined with TMS (e.g., fMRI, EEG, fNIRS) were excluded unless they also had independent MEP amplitude or movement outcome measures. (4) The study had to report original statistics (i.e., not a re-analysis of already published findings or a review) of the intervention on MEP amplitude or movement outcome measures. (5) The study had to be published in English.

2.2. Screening Process

A three-stage screening process was adopted in this study (i.e., title, abstract, and full-text screening). At all screening stages, the two co-first authors independently screened all articles. Conflicts were resolved by discussion between the co-first authors in the title screen phase, with additional input from the final author in the abstract and full text screen stages. First, all 404 titles were screened to evaluate relevance to the current meta-analysis, with 119 titles classified as eligible for further inspection of abstracts (intercoder reliability = 94.55%). Second, the 119 abstracts were assessed for eligibility based on the indicative content provided, with 47 abstracts classified as eligible for further inspection as full texts (intercoder reliability = 96.64%). Third, the full text articles were assessed for eligibility based on the study inclusion criteria, with 34 studies progressing to data extraction (intercoder reliability = 97.87%). The mean intercoder agreement across the three stages of the screening process was 96.35%. Whenever data were missing to compute the effect sizes of interest, this information was requested from the corresponding author for the respective article.

2.3. Data Processing

2.3.1. Data Extraction

The final 34 studies were split between the two co-first authors based on study focus, with author two (ACV) extracting MEP data from 17 studies and author one (SC) extracting movement outcome data from 17 studies, respectively. The co-first authors extracted the primary data (i.e., means, standards deviations, and samples sizes) and necessary methodological information to investigate the proposed moderator effects for the present meta-analysis. Once complete, the final author blind-checked the primary data for all studies and contacted the corresponding authors for studies with missing primary data. If the authors did not respond within one month, or prior to the study data analysis cut-off date (January 2021), the article was excluded from the present meta-analysis. One study that collected MEP data and one study that collected movement outcome data were excluded, leaving 32 studies in the final meta-analysis (see supplementary files for recorded methodological information and primary data for all included articles).

2.3.2. Effect Size Preparation

For both sets of data, Cohen's *d* effect sizes were calculated using the mean, standard deviation, and sample size values for the relevant outcome measures prior to conducting the main data analyses using the METAGEN function from the 'meta' package (Balduzzi et al., 2019) in R studio statistical software (version 2021.09.2 Build 382). The effect size preparation process varied between the studies that collected MEP amplitude data and movement outcome data. All studies that collected MEP amplitude data adopted a repeated measures design, meaning raw data was recorded for all conditions of interest (i.e., AOMI, AO, MI, control). Additional steps were required for studies that collected movement outcome data to account for differences in the movement outcome variables and study designs adopted. When a reduction in a movement outcome measure compared to control was considered an improvement (e.g., reaction time, movement time, mean radial error), the polarity of the calculated effect size was inverted (i.e., positive reversed to negative and vice versa) to ensure an increased value always indicated an improvement.

In order to produce standardized effect sizes that could be compared across the different study designs, effect sizes were calculated for either pre- vs post-test gain comparisons or post- vs post-test comparisons (see supplementary materials for all effect size calculation formulae). Pre-vs-post-test gain comparisons were used to compute effect sizes in studies that adopted mixed experimental designs (i.e., studies that allocated participants to a specific intervention condition and collected data pre- and post-test) as this controls for any pre-test differences that exist between intervention condition groups (Durlak, 2009). Post- vs post-test comparisons were used to compute effect sizes in studies that adopted within-subject (i.e., repeated-measures) or between-subject (i.e., independent intervention condition groups) experimental designs. The number of effect sizes included from each study was not limited as most studies recorded more than one outcome measure for each comparison of interest (i.e., AOMI vs AO, MI, or control). Multiple effect sizes were extracted from studies when MEP amplitudes were collected from more than one target muscle or multiple movement outcome measures were recorded. Cohen's *d* effect size values were converted

into Hedge's *g* effect size values using a small sample size correction formula (Hedges, 1981) for sensitivity analysis.

2.4. Data Analysis

Analyses were conducted using the 'meta' (Balduzzi et al., 2019), 'metafor' (Viechtbauer, 2010) and 'dmetar' (Harrer et al., 2019) packages in R studio (see supplementary materials for R coding used when analyzing the data).

2.4.1. Data Screening

Meta-analyses may be subject to multiple biases (Harrer et al., 2021). To address concerns about publication bias, visual analyses were conducted on the data using funnel plots (Lau et al., 2006), and subsequent statistical analyses were run, if necessary, using Egger's regression (Egger et al., 1997). For any data that showed publication bias, the trim-and-fill procedure (Duval & Tweedie, 2000) was used to determine the number of unpublished studies required to produce an unbiased estimate of the actual effect size. To address concerns about sample size bias, a sensitivity analysis was run by repeating the main analyses with previously calculated Hedge's *g* values. This found no meaningful differences in effect size estimate for the primary comparisons, so Cohen's *d* values are reported in this meta-analysis (cf. Simonsmeier et al., 2020). To address concerns about between-study heterogeneity, outlier diagnostics were completed using the FIND.OUTLIERS function and influence analyses were run using the INFLUENCEANALYSIS function from the 'dmetar' package in R (Harrer et al., 2019). This involved visual inspection of "baujat", "influence" for effect sizes, and "leave-one-out" plots for both effect size and I^2 values. Potential outliers and influential effect sizes were identified across the six primary comparisons for the MEP amplitude and movement outcome data (see supplementary files for detailed overview of data screening results). Removal of the outliers and influential cases had minimal impact on the pooled effect and heterogeneity estimates. Therefore, all effect sizes apart from one effect size for the MI comparison in the movement outcome data (i.e., Marshall et al., 2016: Average distance from the hole) were retained in the meta-analysis to preserve the richness of the data.

2.4.2. Quality Assessment

Study quality was assessed to identify if the studies included in this meta-analysis provide reliably reported data, as well as indicating whether these studies reach an acceptable scientific standard (Borenstein et al., 2009). To address these, the first author subjectively assessed the quality of each study using an assessment scale employed in recent meta-analyses (see e.g., Harris et al., 2021). This quality assessment scale was adapted from the Quality Index (Downs & Black, 1998), the Checklist for the Evaluation of Research Articles (Durant, 1994), and the Appraisal Instrument (Genaidy et al., 2007). The quality assessment checklist and individual scores for each study are provided in the supplementary materials.

2.4.3. Primary and Moderator Effects

The primary meta-analyses for both the MEP amplitude and movement outcome data involved random effects statistical models (Raudenbush, 2009) using the 'meta' package in R (Balduzzi et al., 2019). Overall, ten moderators were chosen across both datasets based on previous meta-analyses and reviews focused on the effects of MI and AO on corticospinal excitability (e.g., Grosprêtre et al., 2016; Naish et al., 2014) and motor skill performance (e.g., Ashford et al., 2006; Simonsmeier et al., 2020, Toth et al., 2020). Four moderators were shared across the MEP amplitude and movement outcome data (*action observation perspective, skill classification, guided attentional focus,* and *kinesthetic imagery ability*). One moderator was specific to the MEP amplitude data (*transcranial magnetic stimulation parameters; point of TMS stimulation, TMS stimulation intensity, number of TMS trials*). Five moderators were specific to the movement outcome data (*population type, physical practice, incorporation of PETTLEP principles, context,* and *intervention volume*). Subgroup analyses and meta regressions were used to examine if these moderators influenced the effects of AOMI on the MEP amplitude and movement outcome data compared to aggregate data from the AO, MI, and control groups. Subgroup analyses were used to compare the effects of AOMI on MEP amplitude and movement outcome data for moderators that permitted sub-division of the primary data sets based on nominal data (i.e., *population type, action observation perspective, skill*

classification, guided attentional focus, point of TMS stimulation, physical practice, incorporation of PETTLEP principles, and context). Regression analyses were used to assess if moderators that consisted of interval data (i.e., *imagery ability, TMS stimulation intensity, number of TMS trials, and intervention volume*) predicted the effects of AOMI on MEP amplitude and movement outcome data.

2.4.3.1. Moderators for Both Sets of Data.

2.4.3.1.1. Action Observation Perspective. Studies have used first-person perspective AO (e.g., Bruton et al., 2020; Romano-Smith et al., 2019) and third-person perspective AO (e.g., Taube et al., 2014; Wright, Wood et al., 2018a) to examine the effects of AOMI on MEP amplitude and movement outcome data. Sub-group analysis was used to compare the effects of AOMI on MEP amplitude and movement outcome data between these two AO perspectives. First-person perspective AO involved the participant viewing the action as if they were performing it (i.e., through their own eyes) and third-person perspective AO involved the participant viewing the action as if another person video recorded them or another person was performing the action (i.e., filmed from a vantage point away from the body). This was determined by checking written text and visual stimuli included in the article.

2.4.3.1.2. Skill Classification. Diverse motor tasks ranging from finger movements (e.g., Meers et al., 2020) to walking (e.g., Kaneko et al., 2018) have been used in previous AOMI literature. Sub-group analysis was used to compare the effects of AOMI on MEP amplitude and movement outcome data for fine vs gross and continuous vs discrete motor tasks. The target movement presented in the AOMI stimuli was classified using a one-dimensional skill classification approach (Spittle, 2021, p.23). Based on this approach, intricate and precise movements using smaller muscle groups (e.g., finger movements) were classed as fine motor tasks; larger muscle movements typically based on fundamental movement patterns (e.g., balance tasks) were classed as gross motor tasks; repetitive movements that have no distinct beginning or end (e.g., walking) were classed as continuous motor tasks; and movements that have an identifiable beginning and end (e.g., putting a

golf ball) were classed as discrete motor tasks. Other skill classification comparisons (e.g., open vs closed skills) were not considered in this moderator category if both motor skill categories were not covered in the synthesized literature.

2.4.3.1.3. Guided Attentional Focus. Studies on AO have demonstrated different effects on movement outcomes (e.g., D'Innocenzo et al., 2016) and MEP amplitudes (e.g., Wright, Wood et al., 2018b) when visual attention is directed, or not, to a specific component of the observed action. More recently, Bruton et al. (2020) showed that allocation of visual attention modulates the effects of AOMI on MEP amplitudes for a finger movement task. Sub-group analysis was used to compare the effects of AOMI on MEP amplitude and movement outcome data for guided attentional focus (i.e., use of instructions to direct attention towards a specific aspect of the observed movement) vs unguided attentional focus (i.e., no such instructions). Studies that did not explicitly state if visual attention was manipulated during AOMI were included in the unguided attentional focus sub-group.

2.4.3.1.4. Kinesthetic Imagery Ability. Kinesthetic imagery is the imagery modality instructed during AOMI and the effects of AOMI on movement outcomes reportedly vary as a function of kinesthetic imagery ability (McNeill et al., 2020). Moderator analysis was used to assess if there was a relationship between kinesthetic imagery ability score and the effects of AOMI on MEP amplitude and movement outcome data. Kinesthetic imagery ability data recorded using valid self-report psychometric scales was extracted from the studies and standardized by reverse-scoring any measures that adopted an inverse scoring system such that higher numbers meant better imagery ability, before converting all scores to percentage values based on the highest possible value.

2.4.3.2. Moderators for MEP Amplitude Data.

2.4.3.2.1. Transcranial Magnetic Stimulation Parameters. AO and MI cause phase-specific changes in corticospinal excitability (see e.g., Grosprêtre et al., 2016; Naish et al., 2014 for reviews) and the number and intensity of TMS trials impacts the reliability of the MEP's evoked during single-pulse TMS (Gangitano et al., 2001; Pellegrini et al., 2018). Sub-group analysis was used to compare the effects of AOMI on MEP amplitude data for TMS delivered at a random point after movement

onset against TMS delivered at a targeted point after movement onset (e.g., at the point of maximum movement of the limb). Timing of TMS stimulation refers to the point at which the single-pulse is delivered based on the movement displayed in the visual stimuli during AOMI. Moderator analyses were used to assess if there was a relationship between number and intensity of TMS trials and the effects of AOMI on MEP amplitude data. The number of TMS stimulations refers to the number of trials where single-pulse TMS is applied to the participant during AOMI. TMS stimulation intensity refers to the relative intensity of the TMS stimulator output applied to the participant during AOMI.

2.4.3.3. Moderators for Movement Outcome Data.

2.4.3.3.1. Population Type. AOMI interventions have been shown to benefit movement outcomes in both neurotypical and neurodivergent populations (e.g., Scott et al., 2020). Sub-group analysis was used to compare the effects of AOMI on movement outcome data for these two population types. Neurotypical populations included individuals who are not characterized by neurologically atypical patterns, thoughts, behavior, or diagnoses, and neurodivergent populations included individuals whose neurological development and state are considered atypical.

2.4.3.3.2. Physical Practice. Studies have explored the effects of AOMI interventions on movement outcomes with (e.g., Marshall & Wright, 2016) and without (e.g., Taube et al., 2014) physical practice. Sub-group analysis was used to compare the effects of AOMI on movement outcome data when used with physical practice vs without physical practice.

2.4.3.3.3. Incorporation of PETTLEP Principles. Some studies have adhered to PETTLEP principles (Holmes & Collins, 2001) when developing and delivering AOMI interventions (e.g., Romano-Smith et al., 2019). Sub-group analysis was used to compare the effects of AOMI on movement outcome data with the inclusion of PETTLEP principles vs without inclusion of PETTLEP principles.

2.4.3.3.4. Context. AOMI interventions have been used to target changes in movement outcomes in sport (e.g., Romano-Smith et al., 2018) and rehabilitation (e.g., Marusic et al., 2018)

contexts. Sub-group analysis compared the effects of AOMI on movement outcome data for sport vs rehabilitation vs other contexts. Studies were classified as sport- or rehabilitation-focused. Studies including movements that did not clearly fall into sports or rehabilitation contexts (e.g., finger movements, ball rotation tasks) were classified as 'other'.

2.4.3.3.5. Intervention Volume. Studies have delivered AOMI interventions over short- (e.g., Bek et al., 2019) and longer-term (e.g., Shimada et al., 2019) periods when investigating their effects on movement outcomes. Moderator analysis was used to assess the relationship between intervention volume (total minutes) and the effects of AOMI on movement outcome data.

3. Results

3.1. Study Characteristics

Overall, this meta-analysis analyzed 113 effect sizes from 32 studies. Of these, 54 effect sizes were from MEP amplitude data (*n* = 16 studies) and 59 effect sizes were from movement outcome data (*n* = 16 studies). Studies included in this meta-analysis were published between 2009 and 2021, with a total sample size of 832 participants split across studies that collected MEP amplitude data (*n* = 234, 77 females, 92 males, 65 undisclosed) and studies that collected movement outcome data (*n* = 589, 281 females, 308 males). The mean age of participants was 27.07 ± 13.48 years and 30.89 ± 20.24 for the two study types, respectively. Some of the articles included in this meta-analysis reported multiple effect sizes, which were falsely analyzed as independent observations (cf. Toth et al., 2020).

3.2. Study Quality

The study quality assessment indicated that all studies included in the meta-analysis displayed a high degree of rigor. For studies that collected MEP amplitude data, the quality assessment scores ranged from 18.75-100%, with a mean of 89.58 ± 22.99%. The most poorly addressed items were 'providing details of a priori sample size determination' and 'consistently reporting effect sizes' with 18.75% and 43.75% of studies satisfying these criteria, respectively (Figure 2a). For studies that collected movement outcome data, quality assessment scores ranged

from 31.25-100%, with a mean of 92.36 ± 17.36%. The most poorly addressed items were 'providing details of a priori sample size determination' and 'applicability of study results to other relevant populations' with 31.25% and 68.75% of studies satisfying these respective criteria (Figure 2b).

To assess whether studies with smaller samples sizes or lower quality studies were likely to bias the results, Pearson's correlation analyses were run between quality assessment scores and effect size, and sample size and effect size using the '*Hmisc*' package in R. For studies that collected MEP amplitude data, there was a moderate negative correlation between quality assessment score and effect size, $r(27) = -0.55$, $p = .003$, and no correlation between sample size and effect size, $r(27) = -0.19$, $p = .34$. For studies that collected movement outcome data, there was no correlation between quality assessment score and effect size, $r(39) = 0.12$, $p = .46$, and a small negative correlation between sample size and effect size, $r(39) = -0.34$, $p = .04$. The mixed correlations between quality assessment score, sample size and effect size indicate a low risk of bias for the studies included in this meta-analysis (*cf*. Harris et al., 2021).

[INSERT FIGURE 2 HERE]

3.3. Primary Effects

3.3.1. MEP Amplitudes

3.3.1.1. AOMI vs Control. Nineteen effect sizes from thirteen studies were used in this analysis (Figure 3). AOMI had a medium positive effect on MEP amplitudes compared to control conditions ($d = 0.54$, 95% CI [0.42, 0.66], $p < .001$). The between-study heterogeneity variance was estimated at $\tau^2 = 0.02$ (95% CI [0.15, 0.67]), with an I^2 value of 0.00% (95% CI [0.00, 48.90]). The prediction interval ranged from $d = -0.12 – 1.21$, indicating that varied intervention effects would be expected for future studies, with these ranging in direction and magnitude from small negative to large positive effects on MEP amplitudes.

[INSERT FIGURE 3 HERE]

3.3.1.2. AOMI vs AO. Twenty-three effect sizes from thirteen studies were used in this analysis (Figure 5). AOMI had a medium positive effect on MEP amplitudes compared to AO

conditions (d = 0.54, 95% CI [0.37, 0.72], p < .001). The between-study heterogeneity variance was estimated at τ^2 0.10 (95% CI [0.00, 0.18]), with an I^2 value of 0.00% (95% CI [0.00, 45.40]). The prediction interval ranged from d = -0.12 – 1.21, indicating that varied intervention effects would be expected for future studies, with these ranging in magnitude from small negative effects to large positive effects on MEP amplitudes.

[INSERT FIGURE 4 HERE]

3.3.1.3. AOMI vs MI. Twelve effect sizes from six studies were used in this analysis (Figure 7). AOMI had a medium positive effect on MEP amplitudes compared to MI conditions (d = 0.55, 95% CI [0.26, 0.84], p = .001). The between-study heterogeneity variance was estimated at τ^2= 0.11 (95% CI [0.00, 0.42]), with an I^2 value of 13.60% (95% CI [0.00, 53.40]). The prediction interval ranged from d = -0.24 – 1.34, indicating that varied intervention effects would be expected for future studies, with these ranging in direction and magnitude from small negative to large positive effects on MEP amplitudes.

[INSERT FIGURE 5 HERE]

3.3.2. Movement Outcomes

3.3.2.1. AOMI vs Control. Twenty-seven effect sizes from twelve studies were used in this analysis (Figure 4). AOMI had a small positive effect on movement outcomes compared to control conditions (d = 0.35, 95% CI [0.07, 0.62], p = .02). The between-study heterogeneity variance was estimated at τ^2 = 0.45 (95% CI [0.22, 0.92), with an I^2 value of 82.60% (95% CI [75.60, 87.60]). The prediction interval ranged from d = -1.07 – 1.76, indicating that varied intervention effects would be expected for future studies, with these ranging in direction from large negative to large positive effects on movement outcomes.

[INSERT FIGURE 6 HERE]

3.3.2.2. AOMI vs AO. Twenty-one effect sizes from nine studies were used in this analysis (Figure 6). AOMI had no significant effect on movement outcomes compared to AO conditions (d = 0.12, 95% CI [-0.17, 0.41], p = .41). The between-study heterogeneity variance was estimated at τ^2=

0.37 (95% CI [0.16, 0.87]), with an I^2 value of 85.10% (95% CI [78.50, 89.70]). The prediction interval ranged from *d* = -1.20 – 1.43, indicating that varied intervention effects would be expected for future studies, with these ranging in direction from large negative to large positive effects on movement outcomes.

[INSERT FIGURE 7 HERE]

3.3.2.3. AOMI vs MI. Eleven effect sizes from six studies were used in this analysis (Figure 8). AOMI had no significant effect on movement outcomes compared to MI conditions (*d* = 0.39, 95% CI [-0.50, 1.28], *p* = .35). The between-study heterogeneity variance was estimated at τ^2 = 1.65 (95% CI [0.71, 5.48]), with an I^2 value of 92.30% (95% CI [88.20, 95.00]). The prediction interval ranged from *d* = -1.02 – 1.79, indicating that varied intervention effects would be expected for future studies, with these ranging in direction and magnitude from large negative to large positive effects on movement outcomes.

[INSERT FIGURE 8 HERE]

3.4. Moderator Effects

The MEP amplitude data (Tables 1 and 2) had two significant moderators and the movement outcome data (Tables 1 and 3) had one significant moderator (see supplementary materials for detailed reporting of all moderator effects). The effects of AOMI on MEP amplitudes differed as a function of TMS stimulation timing (*p* = .01). Specifically, MEP amplitudes were increased to a greater extent during AOMI compared to AO, MI and control conditions when TMS pulses were delivered at random time points (*d* = 0.71, *p* < .001) compared to when they were delivered at targeted timepoints after movement onset (*d* = 0.46, *p* < .001). The effect of AOMI on MEP amplitudes was also positively predicted by the proportionate intensity of the applied TMS single-pulse, *β* = 0.02, *F* [1, 47] = 4.88, *p* = .03. The effect of AOMI on movement outcomes was negatively predicted by the kinesthetic imagery ability score, *β* = - 0.03, *F* [1, 19] = 4.96, *p* = .04. None of the other moderator comparisons reached significance, demonstrating a robust effect of AOMI on both sets of data irrespective of population type, AO perspective, skill classification, and attentional focus

strategy. Additionally, there was a robust effect for the MEP amplitude data irrespective of number of trials and kinesthetic imagery ability, and movement outcome data irrespective of practice incorporation, adherence to PETTLEP principles, context, and intervention volume.

[INSERT TABLE 1 HERE]

[INSERT TABLE 2 HERE]

[INSERT TABLE 3 HERE]

3.5. Publication Bias

Based on the funnel plots of effect size level (see supplementary files) and the results of Egger's regression analyses, publication bias was identified to be unlikely for the control comparison in the MEP amplitude data and for the AO and MI comparisons in the movement outcome data. For the AO comparison in the MEP amplitude data, both the funnel plot of effect size level (Figure 9A, left) and the Egger's regression for random effects indicated asymmetry ($p < .01$). Trim-and-fill analysis proposed 4 missing values (Figure 9A, right), but the effect size changed from a medium (d = 0.54) to a small to medium positive effect (d = 0.43), suggesting minimal effects of publication bias for this dataset. For the MI comparison in the MEP amplitude data, the funnel plot of effect size level (Figure 9B, left) and the Egger's regression for random effects indicated asymmetry ($p < .01$). Trim-and-fill analysis proposed 4 missing values (Figure 9B, right), and the effect size changed from a medium (d = 0.55) to a small positive effect (d = 0.33), suggesting an effect of publication bias for this dataset. For the control comparison in the movement outcome data, both the funnel plot of effect size level (Figure 9C, left) and the Egger's regression for random effects indicated asymmetry (p = .01). Trim-and-fill analysis proposed 5 missing values (Figure 9C, right), and the effect size changed from a small to medium (d = 0.35) to a small positive effect (d = 0.11), suggesting an effect of publication bias for this dataset.

4. Discussion

Since the early reviews introducing AOMI (e.g., Eaves, Riach et al., 2016; Vogt et al., 2013), researchers have studied its effects as a motor simulation intervention. The purpose of the current

study was to methodically collate and synthesize the available MEP amplitude data as an indicator of corticospinal excitability during AOMI engagement, and movement outcome data to assess changes in behavior that result from AOMI interventions. Our primary aim was to establish the effectiveness of AOMI by comparing its effects on MEP amplitudes and movement outcomes against those for AO, MI, and control conditions. Based on previous literature (see Eaves, Riach et al., 2016), it was hypothesized that AOMI would have a small positive effect compared to independent AO or MI, and a moderate positive effect compared to control conditions, for both outcome variables. The results of this meta-analysis partially support this hypothesis. For the MEP amplitude data, a moderate positive effect was reported, indicating increased corticospinal excitability for AOMI compared to AO, MI, and control conditions. For the movement outcome data, AOMI had a small to moderate positive effect compared to control conditions, with no effect reported compared to AO and MI conditions, suggesting AOMI improves behavior, but not to a greater extent than these well-established motor simulation approaches.

4.1. MEP Amplitudes

In this meta-analysis, AOMI had a moderate positive effect on MEP amplitude data, as an indicator of corticospinal excitability, compared to AO, MI and control conditions. TMS studies have consistently reported increased corticospinal facilitation for AOMI across diverse motor tasks such as simple finger movements (Bruton et al., 2020), walking (Kaneko et al., 2018), and basketball free throws (Wright, Wood et al., 2018a). From a theoretical standpoint, the finding that corticospinal excitability was facilitated for AOMI compared to all conditions aligns with the sentiments of the DASH (Eaves et al., 2012, 2014) but may slightly oppose the propositions that AOMI is driven by MI, as made by the VGH (Meers et al., 2020). The DASH proposes that concurrent representations of observed and imagined actions can be maintained as two quasi-encapsulated sensorimotor streams and that these will merge, rather than compete, when a person is overtly and covertly simulating the same action during AOMI. The merging of these two sensorimotor streams is likely to produce more widespread activity in the premotor cortex (see Filimon et al., 2015) than the AO, MI and control

conditions, contributing to increased corticospinal excitability via cortico-cortical connections linking premotor and motor cortices (Fadiga et al., 2005). Alternatively, the VGH suggests that observed and imagined actions are not co-represented, and MI is the driver for increases in motor activity during AOMI. Specifically, Meers et al. (2020) suggest the AO component acts as a visual primer, facilitating the production of more vivid images during AOMI compared to MI conditions. If the propositions of the VGH were true, the increased imagery vividness for AOMI vs MI could be expected to be represented by a smaller difference in MEP amplitude between these two conditions, especially given the participants synthesized in this meta-analysis self-reported being able to generate clear and vivid kinesthetic imagery without a visual primer (mean normalized kinesthetic imagery ability score = 66.06%). Current evidence is conflicting for the DASH and VGH accounts of AOMI (see Bruton et al., 2020; Meers et al., 2020), but both hypotheses offer feasible explanations for the impact of AOMI on the motor system and thus warrant further systematic investigation.

AOMI had a larger effect on MEP amplitudes when single-pulse TMS was delivered randomly during the movement. Changes in corticospinal excitability for AO and MI are reported to mirror the time specificity of actual movements (Gangitano et al., 2001), meaning targeted stimulation at the point of maximum muscle contraction for the observed or imagined action should lead to greatest facilitation in corticospinal excitability (see e.g., Grosprêtre et al., 2016; Naish et al., 2014 for reviews). It is possible that greater excitation for random delivery of single-pulse TMS during AOMI may reflect a discord in timing for AO and MI when synchronized. Specifically, due to the likelihood that individuals will not be able synchronize AO and MI perfectly when performing AOMI (Frank et al., 2020), the use of different stimulation timings for AOMI of an action may be more likely to pick-up the 'optimal timing' for both AO and MI of the action and reduce the predictability of the stimulation and subsequent anticipatory responses of the participants (Loporto et al., 2012).

Additionally, TMS stimulation intensity held a positive correlation with MEP amplitudes during AOMI. When TMS is delivered at rest, MEP amplitudes are less variable at higher stimulation intensities (120-135%) of an individual's resting motor threshold (Pellegrini et al., 2018). However,

538 studies have demonstrated that AO facilitates MEP amplitudes at low- rather than high-intensity
539 (e.g., Loporto et al., 2013). The studies included in this meta-analysis had a mean average
540 stimulation intensity of 116.05 ± 6.95% resting motor threshold. Given the influence of these
541 moderators on MEP amplitudes during AOMI, it is important for future research to explore the time-
542 specific modulation of corticospinal excitability during AOMI and test the impact of different TMS
543 stimulation intensities to increase methodological rigor when investigating the effects of this
544 intervention on corticospinal excitability.

545 **4.2. Movement Outcomes**

546 AOMI had a small to moderate positive effect on movement outcomes compared to control
547 conditions. Such positive effects are evidenced across most studies included in this meta-analysis,
548 with movements ranging from dart throwing (Romano-Smith et al., 2018, 2019) to whole-body
549 balance tasks (Taube et al., 2014), in both neurotypical (e.g., Aoyama et al., 2020) and
550 neurodivergent populations (e.g., Marshall et al., 2020). The increased motor activity during AOMI,
551 as discussed in the previous section, is a possible neurophysiological mechanism for this effect on
552 movement outcomes. Repeated engagement in AOMI, and thus activation of the motor system, has
553 the potential to support repetitive Hebbian modulation of intracortical and subcortical excitatory
554 mechanisms through synaptic plasticity, in a similar manner to physical practice (Holmes & Calmels,
555 2008).

556 From a cognitive perspective, AO and MI help develop mental representations that comprise
557 cognitive information relating to movement execution (Frank et al., 2020). When executing a motor
558 task, a person recalls the relevant mental representation and uses this to guide their movement
559 (Frank et al., 2020). AO and MI may contribute differently to the development of such mental
560 representations, with AO providing sequential and timing information and MI providing sensory
561 information related to the movement (Kim et al., 2017). It is possible combining the two forms of
562 motor simulation during AOMI allows for the effective development of mental representations of

action in the long-term memory, benefitting the physical execution of a motor task (Frank et al., 2020; Kim et al., 2017; Wright et al., 2021).

In contrast, AOMI had no effect on movement outcomes compared to AO and MI conditions. Robust evidence supports the efficacy of AO and MI as interventions to improve motor performance across settings (e.g., AO: Buccino, 2014; MI: Guillot & Collet, 2008). This null finding shows a discrepancy between the effects of AOMI on corticospinal excitability and movement outcomes when compared to AO and MI in this meta-analysis. Such disparity may be in-part explained by the concept of 'neural efficiency'. Experts consistently demonstrate more spatially localized and less intense brain activity during sensorimotor task performance when compared to less-skilled individuals (Neubauer & Fink, 2009). However, Mizuguchi and Kanosue (2017) propose that brain activity is increased during AO, MI, and physical execution of sensorimotor tasks during initial stages of motor learning, with activity in motor regions of the brain reducing for physical execution and increasing for AO and MI of the task as learning progresses. The increased corticospinal excitability during AOMI may therefore only reflect the conscious and focused attentional processes associated with movement execution, rather than the more autonomous and targeted brain activity that underpins skilled execution (Holmes & Wright, 2017).

The findings of this meta-analysis promote AOMI as an effective alternative intervention to AO and MI as well-established approaches, but do not indicate that combining AO and MI simultaneously has an additive benefit towards motor performance. It is worthwhile noting that AOMI had a small to moderate positive effect on movement outcomes compared to these interventions (d = 0.12 and 0.39 for AO and MI, respectively) despite the lack of significant differences. This is an important consideration in applied settings, such as sport and neurorehabilitation, where marginal improvements in motor performance can have practical significance (Lakens, 2013). AOMI interventions are a suitable alternative to AO and MI interventions as this combined approach can address the reported limitations of using either simulation technique in isolation. The capacity to generate and maintain mental images, termed 'imagery ability', is a

complex cognitive process that is variable within- and between-populations (Cumming & Eaves, 2018). Individuals with low imagery ability typically find it difficult to generate and control imagined content during MI interventions, an issue that is not present for AO interventions as specific movement content can be displayed via video (Holmes & Calmels, 2008). However, the effectiveness of AO interventions is dependent on the observer's ability to attend to the most important aspects of the motor task being performed (D'Innocenzo et al., 2016). Based on current recommendations for delivering AOMI interventions (see Wright et al., 2021), AOMI has the capacity to control the visual information displayed via AO whilst directing the individual's attention by getting them to focus on kinesthetic aspects of the movement, subsequently reducing the complexity of MI.

Kinesthetic imagery ability negatively predicted the effects of AOMI on movement outcomes (i.e., the lower the imagery ability the greater the effect AOMI had on movement outcomes). This result opposes McNeill and colleagues' (2020) findings that skilled golfers with good kinesthetic imagery ability improve their golf putting performance after an AOMI intervention (where AO and MI were alternated), but 'poor imagers' do not experience such benefits. It is possible that simultaneous AOMI, where AO and MI are in synchrony, negates the importance of imagery ability for motor skill performance benefits. This is likely due to the individual using the visual stimuli of someone performing the action successfully to facilitate their kinesthetic imagery of the movement, as per the VGH (Meers et al., 2020). This is further supported by our results showing kinesthetic imagery ability did not moderate the effects of AOMI on corticospinal excitability. This result aligns with recent evidence showing no relationship between imagery ability characteristics and corticospinal excitability for novices during AOMI of a complex task (Moriuchi et al., 2020). The discord between the moderating effects of kinesthetic imagery ability for movement outcomes (a negative prediction) and MEP amplitudes (no prediction) may be explained by expertise-based differences in the role of imagery ability during AOMI. For MI, studies have shown that imagery ability is positively correlated with corticospinal excitability for experts but not novices (e.g., Fourkas et al., 2008), and that brain activation patterns vary based on experience in the task being imagined

(Kraeutner et al., 2018). The studies included in this meta-analysis used simple tasks that fell within the motor repertoire of the study sample when recording MEP amplitudes during AOMI, whereas the studies examining movement outcomes typically recruited novice or impaired populations that have not mastered the task being learned/refined via AOMI. It is recommended that future research consider interactions between prior experience/mastery and kinesthetic imagery ability when exploring the effects of AOMI on corticospinal excitability and movement outcomes.

4.4. Limitations and Future Research Recommendations

4.4.1. Neurophysiological Modality

This meta-analysis synthesized MEP amplitude data from AOMI studies using single-pulse TMS. TMS has been widely used as a neurophysiological modality with AOMI as it permits the recording of muscle-specific facilitation in corticospinal excitability, an effect that has been demonstrated robustly for AO and MI (Naish et al., 2014; Grosprêtre et al., 2016). However, studies have shown that AOMI leads to activity in brain regions that would not be activated directly during the delivery of TMS to the primary motor cortex (e.g., rostral prefrontal cortex; Eaves, Behmer, et al., 2016), and therefore not considered within this meta-analysis. Consequently, there is a need to collate and synthesize data on the precise anatomical substrates involved in AOMI using neuroscientific methods with increased spatial resolution. Hardwick et al. (2018) recently performed a large-scale activation likelihood estimation meta-analysis on fMRI data for AO, MI, and movement execution to identify distinct and shared neural regions for these three action states. This approach could be adopted once additional fMRI data is avaliable on AOMI to advance understanding of the neural mechanisms underpinning engagement in this form of simulation.

4.4.2. Study Designs

To-date, AOMI studies have almost entirely explored the short- term effects of this intervention on movement outcomes, using a between-groups comparison at one time point or adopting a pre- vs post-test study design. Whilst this approach is typical in randomized controlled trials of interventions, this does not permit accurate assessment of the long-term changes that result

from AOMI engagement. The benefits of AO and MI on movement outcomes are reportedly greatest during or immediately after training, with the positive outcomes gradually disappearing in the absence of simulated practice (Stevens et al., 2003; Zhang et al., 2019). However, the performance benefits of MI training are retained beyond the intervention period (Simonsmeier et al., 2020), with repetitive engagement in MI inducing neural plasticity during recovery phases when this technique is used to acquire a skill (Ruffino et al., 2017). It remains unclear if the improvements in movement outcome associated with AOMI are maintained after the intervention is withdrawn. Future studies should draw from the methodological approaches adopted in motor learning literature (e.g., Krakauer et al., 2019) to comprehensively examine the effectiveness of AOMI when learning and improving movement outcomes for different populations and motor tasks.

4.4.3. Brain-Behavior Interactions

This meta-analysis synthesized data on MEP amplitudes during AOMI engagement and changes in movement outcomes after AOMI interventions. The crossover between these effects could not be analyzed in the current meta-analysis as few studies have collected both measures for AOMI. The results of this meta-analysis showed that AOMI had a positive effect on MEP amplitudes compared to control, AO, and MI conditions, but these positive effects were only present compared to control conditions for movement outcomes. Collecting neurophysiological responses, such as MEP amplitudes, alongside movement outcome measures will advance our understanding of the relationship between the changes in brain activity and motor skill performance for this intervention. Such integration of measures has shown that corticospinal excitability during MI is related to the magnitude of motor cortical adaptations after MI training (Yoxon & Welsh, 2020), and that MI and physical practice lead to different changes in brain activity and subsequent movement execution after equal training bouts (Kraeutner et al., 2020). Future research should combine neuroscientific and human movement science methods to elucidate the neural mechanisms underlying improvements in motor skill performance and learning through AOMI.

5. Conclusion

This meta-analysis synthesizes the existing MEP amplitude and movement outcome data for AOMI to compare its effectiveness against AO, MI, and control conditions. The results of the present meta-analysis report that AOMI has a moderate positive effect on MEP amplitudes, as an indicator of corticospinal excitability, compared to AO, MI, and control conditions. AOMI also has a small to moderate positive effect on movement outcomes compared to control conditions, with no such effects demonstrated when compared to AO and MI conditions. TMS stimulation timing and intensity moderated the effects of AOMI on MEP amplitude data, whereas kinesthetic imagery ability moderated the effects of AOMI on movement outcome data. Overall, the findings of this meta-analysis support the effectiveness of AOMI as an alternative intervention to AO and MI, two well established interventions, as it addresses the limitations of these approaches in isolation when targeting increased activity in motor regions of the brain and improvements in motor skill performance. A more methodologically diverse approach that integrates brain and behavior is needed in future AOMI research to advance the current state of knowledge for this intervention.

Funding statement: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

6. References

Aoyama, T., Kaneko, F., & Kohno, Y. (2020). Motor imagery combined with action observation training optimized for individual motor skills further improves motor skills close to a plateau. *Human Movement Science*, *73*, 102683. https://doi.org/10.1016/j.humov.2020.102683

Ashford, D., Bennett, S. J., & Davids, K. (2006). Observational modeling effects for movement dynamics and movement outcome measures across differing task constraints: A meta-analysis. *Journal of Motor Behavior*, *38*(3), 185-205. https://doi.org/10.3200/JMBR.38.3.185-205

Balduzzi S, Rücker G, Schwarzer G (2019), How to perform a meta-analysis with R: a practical tutorial, *Evidence-Based Mental Health, 22*(4), 153-160. https://dx.doi.org/10.1136/ebmental-2019-300117

Bek, J., Gowen, E., Vogt, S., Crawford, T. J. & Poliakoff, E. (2019). Combined action observation and motor imagery influences hand movement amplitude in Parkinson's disease. *Parkinsonism and Related Disorders, 61*, 126-131. https://doi.org/10.1016/j.parkreldis.2018.11.001

Berends, H., Wolkorte, R., Ijzerman, M., & van Putten, M. (2013). Differential cortical activation during observation and observation-and-imagination. *Experimental Brain Research*, *229*(3), 337-345. https://doi.org/10.1007/s00221-013-3571-8

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John Wiley & Sons.

Bruton, A., Holmes, P., Eaves, D., Franklin, Z., & Wright, D. (2020). Neurophysiological markers discriminate different forms of motor imagery during action observation. *Cortex*, *124*, 119-136. https://doi.org/10.1016/j.cortex.2019.10.016

Buccino, G. (2014). Action observation treatment: A novel tool in neurorehabilitation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130185. https://doi.org/10.1098/rstb.2013.0185

Caligiore, D., Mustile, M., Spalletta, G., & Baldassarre, G. (2017). Action observation and motor imagery for rehabilitation in Parkinson's disease: A systematic review and an integrative hypothesis. *Neuroscience & Biobehavioral Reviews*, *72*, 210-222. https://doi.org/10.1016/j.neubiorev.2016.11.005

Castro, F., Bryjka, P.A., Di Pino, G., Vuckovic, A., Nowicky, A. & Bishop, D. (2021). *Brain and Cognition*, *152,* 105768. https://doi.org/10.1016/j.bandc.2021.105768

Cumming, J., & Eaves, D. L. (2018). The nature, measurement, and development of imagery ability. *Imagination, Cognition and Personality*, *37*(4), 375-393. https://doi.org/10.1177%2F0276236617752439

D'Innocenzo, G., Gonzalez, C. C., Williams, A. M., & Bishop, D. T. (2016). Looking to learn: the effects of visual guidance on observational learning of the golf swing. *PloS one, 11*(5), e0155442. https://doi.org/10.1371/journal.pone.0155442

De Vries, S., & Mulder, T. (2007). Motor imagery and stroke rehabilitation: A critical discussion. *Journal of Rehabilitation Medicine*, *39*(1), 5-13. https://doi.org10.2340/16501977-0020

Di Rienzo, F., Joassy, P., Kanthack, T., MacIntyre, T., Debarnot, U., & Blache, Y. et al. (2019). Effects of Action Observation and Action Observation Combined with Motor Imagery on Maximal Isometric Strength. *Neuroscience, 418*, 82-95. https://doi.org/10.1016/j.neuroscience.2019.08.025

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*, *52*(6), 377-384. http://dx.doi.org/10.1136/jech.52.6.377

Durant, J. (1994). What is scientific literacy? *European Review*, *2*(1), 83-89. https://doi.org/10.1017/S1062798700000922

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology*, *34*(9), 917-928. https://doi.org/10.1093/jpepsy/jsp004

Duval, S., & Tweedie, R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89-98. https://doi.org/10.1080/01621459.2000.10473905

Eaves, D. L., Behmer, L., & Vogt, S. (2016). EEG and behavioural correlates of different forms of motor imagery during action observation in rhythmical actions. *Brain and Cognition*, *106*, 90-103. https://doi.org/10.1016/j.bandc.2016.04.013

Eaves, D. L., Haythornthwaite, L., & Vogt, S. (2014). Motor imagery during action observation modulates automatic imitation effects in rhythmical actions. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00028

Eaves, D. L., Riach, M., Holmes, P. S., & Wright, D. J. (2016). Motor imagery during action observation: a brief review of evidence, theory and future research opportunities. *Frontiers in Neuroscience*, *10*, 514. https://doi.org/10.3389/fnins.2016.00514

Eaves, D. L., Turgeon, M., & Vogt, S. (2012). Automatic imitation in rhythmical actions: Kinematic fidelity and the effects of compatibility, delay, and visual monitoring. *PLoS One*, *7*(10), e46728. https://doi.org/10.1371/journal.pone.0046728

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629-634. https://doi.org/10.1136/bmj.315.7109.629

Emerson, J. R., Binks, J. A., Scott, M. W., Kenny, R. P., & Eaves, D. L. (2018). Combined action observation and motor imagery therapy: A novel method for post-stroke motor rehabilitation. *AIMS Neuroscience*, *5*(4), 236-252. https://doi.org/10.3934/Neuroscience.2018.4.236

Fadiga, L., Craighero, L., & Olivier, E. (2005). Human motor cortex excitability during the perception of others' action. *Current Opinion in Neurobiology*, *15*(2), 213-218. https://doi.org/10.1016/j.conb.2005.03.013

Filimon, F., Rieth, C., Sereno, M., & Cottrell, G. (2015). Observed, executed, and imagined action representations can be decoded from ventral and dorsal areas. *Cerebral Cortex*, *25*(9), 3144-3158. https://doi.org/10.1093/cercor/bhu110

Fourkas, A. D., Bonavolontà, V., Avenanti, A., & Aglioti, S. M. (2008). Kinesthetic imagery and tool-specific modulation of corticospinal representations in expert tennis players. *Cerebral Cortex*, *18*(10), 2382-2390. https://doi.org/10.1093/cercor/bhn005

Frank, C., Kim, T., & Schack, T. (2018). Observational practice promotes action-related order formation in long-term memory: Investigating action observation and the development of cognitive representation in complex motor action. *Journal of Motor Learning and Development*, *6*(1), 53-72. https://psycnet.apa.org/doi/10.1123/jmld.2017-0007

Frank, C., Wright, D., & Holmes, P. (2020). Mental simulation and neurocognition: Advances for motor imagery and action observation training in sport. In D. Hackfort & R. J. Schinke (Eds.), *Routledge International Encyclopaedia of Sport and Exercise Psychology* (Vol. 2., pp. 411-428). Routledge. https://doi.org/10.4324/9781315187228

Fogassi, M., Gallese, V., Fogassi, L., Gallese, M., Raos, V., Rizzolatti, G., & Rizzolatti, M. (2001). Phase-specific modulation of cortical motor output during movement observation. *NeuroReport*, *12*(7), 1489-1492.

Fowkes, G. R., Lemaitre, A. M., Lemasters, G. K., Lockey, J., Succop, P., Deddens, J., Sobeih, T., & Dunning, K. (2007). An epidemiological appraisal instrument – a tool for evaluation of epidemiological studies. *Ergonomics*, *50*(6), 920-960. https://doi.org/10.1080/00140130701237667

Frank, C., Land, W. M., Schack, T., Mesagno, C., Bloomfield, J., & Lebon, F. (2016). Motor imagery and cortico-spinal excitability: A review. *European Journal of Sport Science*, *16*(3), 317-324. https://doi.org/10.1080/17461391.2015.1024756

Guillot, A., & Collet, C. (2008). Construction of the motor imagery integrative model in sport: A review and theoretical investigation of motor imagery use. *International Review of Sport and Exercise Psychology*, *1*(1), 31-44. https://doi.org/10.1080/17509840701823139

Hardwick, R., Caspers, S., Eickhoff, S., & Swinnen, S. (2018). Neural correlates of action: Comparing meta-analyses of imagery, observation, and execution. *Neuroscience & Biobehavioral Reviews*, *94*, 31-44. https://doi.org/10.1016/j.neubiorev.2018.08.003

Harrer, M., Cuijpers, P., Furukawa, T. & Ebert, D. D. (2019). dmetar: Companion R Package for the Guide 'DoingMeta-Analysis in R'. R package version 0.0.9000. URL: http://dmetar.protectlab.org/.

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing Meta-Analysis with R: A Hands-On Guide.* Chapmann & Hall; CRC Press. URL: https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/

Harris, D. J., Allen, K. L., Vine, S. J., & Wilson, M. R. (2021). A systematic review and meta-analysis of the relationship between flow states and performance. *International Review of Sport and Exercise Psychology*, 1-29. https://doi.org/10.1080/1750984X.2021.1929402

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107-128. https://doi.org/10.3102/10769986006002107

Holmes, P. S., & Calmels, C. (2008). A neuroscientific review of imagery and observation use in sport. *Journal of Motor Behavior*, *40*(5), 433-445. https://doi.org/10.3200/jmbr.40.5.433-445

Holmes, P. S., & Collins, D. J. (2001). The PETTLEP approach to motor imagery: A functional equivalence model for sport psychologists. *Journal of Applied Sport Psychology, 13*(1), 60-83. https://doi.org/10.1080/10413200109339004

Holmes, P. S., & Wright, D. J. (2017). Motor cognition and neuroscience in sport psychology. *Current Opinion in Psychology*, *16*, 43-47. https://doi.org/10.1016/j.copsyc.2017.03.009

Jeannerod, M. (1994). Motor representations and reality. *Behavioral and Brain Sciences*, *17*(2), 229-245. https://doi.org/10.1017/s0140525x0003435x

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, *14*(1), S103-S109. https://doi.org/10.1006/nimg.2001.0832

Jeannerod, M. (2006). The origin of voluntary action. History of a physiological concept. *Comptes Rendus Biologies*, *329*(5-6), 354-362. https://doi.org/10.1016/j.crvi.2006.03.017

Kaneko, N., Masugi, Y., Yokoyama, H., & Nakazawa, K. (2018). Difference in phase modulation of corticospinal excitability during the observation of the action of walking, with and without motor imagery. *NeuroReport*, *29*(3), 169-173. https://doi.org/10.1097/wnr.0000000000000941

Kawasaki, T., Tozawa, R., & Aramaki, H. (2018). Effectiveness of using an unskilled model in action observation combined with motor imagery training for early motor learning in elderly people: A preliminary study. *Somatosensory & Motor Research*, *35*(3-4), 204-211. https://doi.org/10.1080/08990220.2018.1527760

Kim, T., Frank, C., & Schack, T. (2017). A systematic investigation of the effect of action observation training and motor imagery training on the development of mental representation structure and skill performance. *Frontiers in Human Neuroscience*, *11*, 499. https://doi.org/10.3389/fnhum.2017.00499

Kraeutner, S. N., McWhinney, S. R., Solomon, J. P., Dithurbide, L., & Boe, S. G. (2018). Experience modulates motor imagery-based brain activity. *European Journal of Neuroscience*, *47*(10), 1221-1229. https://doi.org/10.1111/ejn.13900

Kraeutner, S. N., Stratas, A., McArthur, J. L., Helmick, C. A., Westwood, D. A., & Boe, S. G. (2020). Neural and behavioral outcomes differ following equivalent bouts of motor imagery or physical practice. *Journal of Cognitive Neuroscience*, *32*(8), 1590-1606. https://doi.org/10.1162/jocn_a_01575

Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., & Haith, A. M. (2019). Motor Learning. *Comprehensive Physiology*, *9*(2), 613–663. https://doi.org/10.1002/cphy.c170043

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lau, J., Ioannidis, J. P., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *British Medical Journal*, *333*(7568), 597-600. https://doi.org/10.1136/bmj.333.7568.597

Loporto, M., Holmes, P. S., Wright, D. J., & McAllister, C. J. (2013). Reflecting on mirror mechanisms: motor resonance effects during action observation only present with low-intensity transcranial magnetic stimulation. *PloS one*, *8*(5), e64911. https://doi.org/10.1371/journal.pone.0064911

Loporto, M., McAllister, C. J., Edwards, M. G., Wright, D. J., & Holmes, P. S. (2012). Prior action execution has no effect on corticospinal facilitation during action observation. *Behavioural Brain Research, 231,* 124e129. https://doi.org/10.1016/j.bbr.2012.03.009.

MacIntyre, T., Moran, A., Collet, C., & Guillot, A. (2013). An emerging paradigm: A strength-based approach to exploring mental imagery. *Frontiers in Human Neuroscience*, *7,* 104. https://doi.org/10.3389/fnhum.2013.00104

Marshall, B., & Wright, D. (2016). Layered stimulus response training versus combined action observation and imagery: Effects on golf putting performance and imagery ability characteristics. *Journal of Imagery Research in Sport and Physical Activity*, *11*(1), 35-46. https://doi.org/10.1515/jirspa-2016-0007

Marshall, B., Wright, D., Holmes, P., & Wood, G. (2019). Combining action observation and motor imagery improves eye-hand coordination during novel visuomotor task performance. *Journal of Motor Behavior*, *52*(3), 333-341. https://doi.org/10.1080/00222895.2019.1626337

Marshall, B., Wright, D., Holmes, P., Williams, J., & Wood, G. (2020). Combined action observation and motor imagery facilitates visuomotor adaptation in children with developmental coordination disorder. *Research in Developmental Disabilities*, *98*, 103570. https://doi.org/10.1016/j.ridd.2019.103570

Marusic, U., Grosprêtre, S., Paravlic, A., Kovač, S., Pišot, R., & Taube, W. (2018). Motor imagery during action observation of locomotor tasks improves rehabilitation outcome in older adults after total hip arthroplasty. *Neural Plasticity, 2018,* 1-9. https://doi.org/10.1155/2018/5651391

McNeill, E., Ramsbottom, N., Toth, A. J., & Campbell, M. J. (2020). Kinaesthetic imagery ability moderates the effect of an AO+ MI intervention on golf putt performance: A pilot study. *Psychology of Sport and Exercise*, *46*, 101610. https://doi.org/10.1016/j.psychsport.2019.101610

Meers, R., Nuttall, H., & Vogt, S. (2020). Motor imagery alone drives corticospinal excitability during concurrent action observation and motor imagery. *Cortex*, *126*, 322-333. https://doi.org/10.1016/j.cortex.2020.01.012

Mizuguchi, N., Kanosue, K., 2017. Changes in brain activity during action observation and motor imagery: their relationship with motor learning. Prog. Brain Res. 234, 189–204. https://doi.org/10.1016/bs.pbr.2017.08.008.

Moriuchi, T., Nakashima, A., Nakamura, J., Anan, K., Nishi, K., Matsuo, T., Hasegawa, T., Mitsunaga, W., Iso, N., & Higashi, T. (2020). The vividness of motor imagery is correlated with corticospinal excitability during combined motor imagery and action observation. *Frontiers in Human Neuroscience*, *14*, 581652. https://doi.org/10.3389/fnhum.2020.581652

Mouthon, A., Ruffieux, J., Keller, M., & Taube, W. (2016). Age-related differences in corticospinal excitability during observation and motor imagery of balance tasks. *Frontiers in Aging Neuroscience*, *8,* 317. https://doi.org/10.3389/fnagi.2016.00317

Mouthon, A., Ruffieux, J., Wälchli, M., Keller, M. & Taube, W. (2015). Task-dependent changes of corticospinal excitability during observation and motor imagery of balance tasks. *Neuroscience, 10*(303), 535-543. https://doi.org/10.1016/j.neuroscience.2015.07.031

Naish, K., Houston-Price, C., Bremner, A., & Holmes, N. (2014). Effects of action observation on corticospinal excitability: Muscle specificity, direction, and timing of the mirror response. *Neuropsychologia*, *64*, 331-348. https://doi.org/10.1016/j.neuropsychologia.2014.09.034

Nedelko, V., Hassa, T., Hamzei, F., Schoenfeld, M., & Dettmers, C. (2012). Action imagery combined with action observation activates more corticomotor regions than action observation alone. *Journal of Neurologic Physical Therapy*, *36*(4), 182-188. https://doi.org/10.1097/npt.0b013e318272cad1

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. *Neuroscience & Biobehavioral Reviews*, *33*(7), 1004-1023. https://doi.org/10.1016/j.neubiorev.2009.04.001

Neuman, B., & Gray, R. (2013). A direct comparison of the effects of imagery and action observation on hitting performance. *Movement & Sport Sciences - Science & Motricité*, (79), 11-21. https://doi.org/10.1051/sm/2012034

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *British Medical Journal,* 372. https://doi.org/10.1136/bmj.n160

Pellegrini, M., Zoghi, M., & Jaberzadeh, S. (2018). The effect of transcranial magnetic stimulation test intensity on the amplitude, variability and reliability of motor evoked potentials. *Brain Research*, *1700*, 190-198. https://doi.org/10.1016/j.brainres.2018.09.002

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The Handbook of Research Synthesis and Meta-Analysis*, *2*, 295-316. Russel Sage Foundation.

Romano Smith, S., Wood, G., Coyles, G., Roberts, J., & Wakefield, C. (2019). The effect of action observation and motor imagery combinations on upper limb kinematics and EMG during dart-throwing. *Scandinavian Journal of Medicine & Science in Sports*, *29*(12), 1917-1929. https://doi.org/10.1111/sms.13534

Romano-Smith, S., Wood, G., Wright, D., & Wakefield, C. (2018). Simultaneous and alternate action observation and motor imagery combinations improve aiming performance. *Psychology of Sport and Exercise*, *38*, 100-106. https://doi.org/10.1016/j.psychsport.2018.06.003

Rothwell, J. C. (1997). Techniques and mechanisms of action of transcranial stimulation of the human motor cortex. *Journal of Neuroscience Methods*, *74*(2), 113-122. https://doi.org/10.1016/S0165-0270(97)02242-5

Rubino, C., Papaxanthis, C., & Lebon, F. (2017). Neural plasticity during motor learning with motor imagery practice: Review and perspectives. *Neuroscience*, *341*, 61-78. https://doi.org/10.1016/j.neuroscience.2016.11.023

Scott, M. W., Wood, G., Holmes, P. S., Williams, J., Marshall, B., & Wright, D. J. (2021). Combined action observation and motor imagery: An intervention to combat the neural and behavioural deficits associated with developmental coordination disorder. *Neuroscience & Biobehavioral Reviews*. *127*, 638-646. https://doi.org/10.1016/j.neubiorev.2021.05.015

Scott, M., Taylor, S., Chesterton, P., Vogt, S., & Eaves, D. (2018). Motor imagery during action observation increases eccentric hamstring force: an acute non-physical intervention. *Disability and Rehabilitation*, *40*(12), 1443-1451. https://doi.org/10.1080/09638288.2017.1300333

Scott, M.W., Emerson, J.R., Dixon, J., Tayler, M. A. & Eaves, D. L. (2019). Motor imagery during action observation enhances automatic imitation in children with and without developmental coordination disorder. *Journal of Experimental Child Psychology, 183*, 242-260. https://doi.org/10.1016/j.jecp.2019.03.001

Shimada, K., Onishi, T., Ogawa, Y., Yamauchi, J., & Kawada, S. (2019). Effects of motor imagery combined with action observation training on the lateral specificity of muscle strength in healthy subjects. *Biomedical Research*, *40*(3), 107-113. https://doi.org/10.2220/biomedres.40.107

Simonsmeier, B., Androniea, M., Buecker, S., & Frank, C. (2020). The effects of imagery interventions in sports: A meta-analysis. *International Review of Sport and Exercise Psychology*, 1-22. https://doi.org/10.1080/1750984x.2020.1780627

Spittle, M. (2021). *Motor learning and skill acquisition: Applications for physical education and sport*. Bloomsbury Publishing.

Ste-Marie, D. M., Lelievre, N., & St. Germain, L. (2020). Revisiting the Applied Model for the Use of Observation: A Review of Articles Spanning 2011–2018. *Research quarterly for exercise and sport*, *91*(4), 594-617.

Ste-Marie, D., Law, B., Rymal, A., Jenny, O., Hall, C., & McCullagh, P. (2012). Observation interventions for motor skill learning and performance: an applied model for the use of observation. *International Review of Sport and Exercise Psychology*, *5*(2), 145-176. https://doi.org/10.1080/1750984x.2012.665076

Stevens, J. A., & Stoykov, M. E. (2003). Using motor imagery in the rehabilitation of hemiparesis. *Archives of Physical Medicine and Rehabilitation*, *84*(7), 1090–1092. https://doi.org/10.1016/s0003-9993(03)00042-x

Taube, W., Lorch, M., Zeiter, S. & Keller, M. (2014). Non-physical practice improves task performance in an unstable, perturbed environment: Motor imagery and observational balance training. *Frontiers in Human Neuroscience, 8,* 972. https://doi.org/10.3389/fnhum.2014.00972

Taube, W., Mouthon, M., Leukel, C., Hoogewoud, H., Annoni, J., & Keller, M. (2015). Brain activity during observation and motor imagery of different balance tasks: An fMRI study. *Cortex*, *64*, 102-114. https://doi.org/10.1016/j.cortex.2014.09.022

Toth, A., McNeill, E., Hayes, K., Moran, A., & Campbell, M. (2020). Does mental practice still enhance performance? A 24 Year follow-up and meta-analytic replication and extension. *Psychology of Sport and Exercise*, *48*, 101672. https://doi.org/10.1016/j.psychsport.2020.101672

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. https://doi.org/10.18637/jss.v036.i03

Villiger, M., Estévez, N., Hepp-Reymond, M., Kiper, D., Kollias, S., Eng, K., & Hotz-Boendermaker, S. (2013). Enhanced activation of motor execution networks using action observation combined with imagination of lower limb movements. *PLoS One*, *8*(8), e72403. https://doi.org/10.1371/journal.pone.0072403

Vogt, S., Rienzo, F., Collet, C., Collins, A., & Guillot, A. (2013). Multiple roles of motor imagery during action observation. *Frontiers in Human Neuroscience*, *7,* 807. https://doi.org/10.3389/fnhum.2013.00807

Wright D. J., Williams J., Holmes P. S. (2014) Combined action observation and imagery facilitates corticospinal excitability. *Frontiers in Human Neuroscience,* 27(8), 951. https://doi.org/10.3389/fnhum.2014.00951

Wright, D. J., Frank, C., & Bruton, A. M. (2021). Recommendations for combining action observation and motor imagery interventions in sport. *Journal of Sport Psychology in Action,* 1-13. https://doi.org/10.1080/21520704.2021.1971810

Wright, D.J., McCormick, S.A., Williams, J. & Holmes, P.S. (2016). Viewing instructions accompanying action observation modulate corticospinal excitability. *Frontiers in Human Neuroscience*, 1, 10(17). https://doi.org/10.3389/fnhum.2016.00017

Wright, D., Wood, G., Eaves, D., Bruton, A., Frank, C., & Franklin, Z. (2018a). Corticospinal excitability is facilitated by combined action observation and motor imagery of a basketball free throw. *Psychology of Sport and Exercise, 39*, 114-121. https://doi.org/10.1016/j.psychsport.2018.08.006

Wright, D. J., Wood, G., Franklin, Z. C., Marshall, B., Riach, M., & Holmes, P. S. (2018b). Directing visual attention during action observation modulates corticospinal excitability. *Plos One, 13*(1), e0190165. https://doi.org/10.1371/journal.pone.0190165.

Yoxon, E., & Welsh, T. (2020). Motor system activation during motor imagery is positively related to the magnitude of cortical plastic changes following motor imagery training. *Behavioural Brain Research*, *390*, 112685. https://doi.org/10.1016/j.bbr.2020.112685

Zhang B, Kan L, Dong A, Zhang J, Bai Z, Xie Y, et al. (2019) The effects of action observation training on improving upper limb motor functions in people with stroke: A systematic review and meta-analysis. *PLoS One*, *14*(8), e0221166. https://doi.org/10.1371/journal.pone.0221166

7. Figure Captions

Figure 1. Flow chart of the meta-analysis literature search and selection procedures

Figure 2. Bar chart displaying quality assessment items and scores for all studies that collected (A) MEP amplitude data and (B) movement outcome data included in the meta-analysis. The blue bar indicates the number of studies that satisfied, and the red bar indicates the number of studies that did not satisfy, each of the respective quality assessment criteria.

Figure 3. Forest plot of effect sizes (*d*) for all studies included in the meta-analysis that compared MEP amplitude data for AOMI vs control conditions. The combined estimate and 95% confidence interval (black diamond) indicates AOMI has a moderate positive effect on MEP amplitudes compared to control conditions. The size of each grey square indicates the weight of the study effect size in the combined analysis. Multiple effect sizes are reported for a study if it recorded MEP amplitude data from more than one target muscle. The thick black line at the base of the plot represents the prediction interval, which is the range that 95% of future studies are likely to fall within.

Figure 4. Forest plot of effect sizes (*d*) for all studies included in the meta-analysis that compared MEP amplitude data for AOMI vs AO conditions. The combined estimate and 95% confidence interval (black diamond) indicates AOMI has a moderate positive effect on MEP amplitudes compared to AO conditions. The size of each grey square indicates the weight of the study effect size in the combined analysis. Multiple effect sizes are reported for a study if it recorded MEP amplitude data from more than one target muscle. The thick black line at the base of the plot represents the prediction interval, which is the range that 95% of future studies are likely to fall within.

Figure 5. Forest plot of effect sizes (*d*) for all studies included in the meta-analysis that compared MEP amplitude data for AOMI vs MI conditions. The combined estimate and 95% confidence interval (black diamond) indicates AOMI has a moderate positive effect on MEP amplitudes compared to MI conditions. The size of each grey square indicates the weight of the study effect size in the combined analysis. Multiple effect sizes are reported for a study if it recorded MEP amplitude data from more than one target muscle. The thick black line at the base of the plot represents the prediction interval, which is the range that 95% of future studies are likely to fall within.

Figure 6. Forest plot of effect sizes (*d*) for all studies included in the meta-analysis that compared movement outcome data for AOMI vs control conditions. The combined estimate and 95% confidence interval (black diamond) indicates AOMI has a small to moderate positive effect on movement outcomes compared to control conditions. The size of each grey square indicates the weight of the study effect size in the combined analysis. Multiple effect sizes are reported for a study if it recorded more than one movement outcome variable. The thick black line at the base of the plot represents the prediction interval, which is the range that 95% of future studies are likely to fall within.

Figure 7. Forest plot of effect sizes (*d*) for all studies included in the meta-analysis that compared movement outcome data for AOMI vs AO conditions. The combined estimate and 95% confidence interval (black diamond) indicates AOMI has no significant effect on movement outcomes compared to AO conditions. The size of each grey square indicates the weight of the study effect size in the combined analysis. Multiple effect sizes are reported for a study if it recorded more than one movement outcome variable. The thick black line at the base of the plot represents the prediction interval, which is the range that 95% of future studies are likely to fall within.

Figure 8. Forest plot of effect sizes (*d*) for all studies included in the meta-analysis that compared movement outcome data for AOMI vs MI conditions. The combined estimate and 95% confidence interval (black diamond) indicates AOMI has no significant effect on movement outcomes compared to MI conditions. The size of each grey square indicates the weight of the study effect size in the combined analysis. Multiple effect sizes are reported for a study if it recorded more than one

movement outcome variable. The thick black line at the base of the plot represents the prediction interval, which is the range that 95% of future studies are likely to fall within.

Figure 9. Funnel plot of effect sizes (Cohen's *d*) versus standard error before (left) and after (right) performing Duval and Tweedie's Trim-and-fill analysis for the (A) AO comparison in the MEP amplitude data, (B) MI comparison in the MEP amplitude data, and (C) control comparison in the movement outcome data. Grey circles represent existing effects included in the meta-analysis and white circles represent potential unpublished effects. The dashed line funnel plot indicates the pooled effect size estimate (vertical line) and the idealized funnel shape for each comparison. The contour-enhanced funnel plots display the significance of the effects for each comparison. Individual effect sizes falling inside the grey ($p < .05$) and white ($p < .01$) funnel boundaries represent significant effects for each comparison.

8. Tables

Table 1. Moderator analyses for the MEP amplitude and movement outcome data

Moderator	N	K	Beta	P	Sig	*df*	Q	I^2
MEP Amplitude								
Kinesthetic Imagery Ability	6	19	-0.01	**.54**	**No**	17	7.89	0.00
Stimulation Intensity	14	49	0.02	**.03**	**Yes**	47	34.04	0.00
Number of trials	15	53	0.00	**.21**	**No**	51	38.63	0.00
Movement Outcome								
Intervention Volume	14	53	0.00	**.65**	**No**	51	321.27	83.33
Kinesthetic Imagery Ability	5	21	-0.03	**.04**	**Yes**	19	208.29	89.50

Note. N = number of studies, K = number of effect sizes, Beta = regression coefficient, P = significance value, Sig = significance threshold met or not, df = degrees of freedom, Q = measure of variation around average, I^2 *= measure of proportion of observed variance (%).*

1079 **Table 2. Sub-group analyses for the MEP amplitude data**

Subgroup	N	K	SMD	95% CI		P	Sig	Q	I^2
Action Observation Perspective									
First-person	9	23	0.48	0.39	0.58	<.001	Yes	8.91	0.00
Third-person	7	31	0.55	0.40	0.71	<.001	Yes	30.97	3.10
Between						**.44**	**No**	0.59	
Directed Attentional Focus									
Yes	2	2	0.38	-0.51	1.27	<.001	Yes	0.06	0.00
No	15	52	0.52	0.43	0.62	<.001	Yes	39.97	0.00
Between						**.34**	**No**	3.00	
Skill Classification 1									
Fine	10	25	0.48	0.39	0.57	<.001	Yes	9.25	0.00
Gross	6	29	0.57	0.41	0.73	<.001	Yes	30.38	7.80
Between						**.33**	**No**	0.95	
Skill Classification 2									
Discrete	12	33	0.50	0.39	0.60	<.001	Yes	21.43	0.00
Continuous	4	21	0.57	0.38	0.76	<.001	Yes	18.45	0.00
Between						**.49**	**No**	0.49	
TMS Stimulation Timing									
Targeted	12	35	0.46	0.35	0.56	<.001	Yes	11.73	0.00
Random	4	19	0.71	0.53	0.88	<.001	Yes	24.29	0.00
Between						**<.001**	**Yes**	6.41	

1080 *Note. N = number of studies, K = number of effect sizes, SMD = standardized mean difference, 95% CI = lower and upper confidence intervals, P = significance*
1081 *value, Sig = significance threshold met or not, Q = measure of variation around average, I^2 = measure of proportion of observed variance (%).*

1082 **Table 3. Sub-group analyses for the movement outcome data**

Subgroup	N	K	SMD	95% CI		P	Significance	Q	I^2
Action Observation Perspective									
First-person	9	31	0.21	-0.10	0.52	.19	No	301.90	90.10
Third-person	7	28	0.19	-0.05	0.43	.10	No	139.39	80.60
Between						**.94**	**No**	0.01	
Context									
Sport	4	17	0.50	0.05	0.95	.03	Yes	182.47	91.20
Rehabilitation	7	29	0.16	-0.07	0.38	.12	No	159.26	82.40
Other	5	12	0.07	-0.22	0.37	.62	No	67.07	83.60
Between						**.29**	**No**	2.46	
Directed Attentional Focus									
Yes	4	15	0.15	-0.29	0.59	.52	No	89.94	84.40
No	12	43	0.25	0.05	0.45	.01	Yes	316.70	86.70
Between						**.67**	**No**	0.18	
PETTLEP									
Yes	5	20	0.46	0.04	0.88	.03	Yes	183.34	89.60
No	11	38	0.08	-0.11	0.27	.39	No	231.92	84.00
Between						**.11**	**No**	2.57	
Physical Practice									
Yes	3	6	0.80	-0.04	1.63	.06	No	60.26	91.70
No	13	52	0.17	-0.06	0.36	.06	No	349.48	85.40
Between						**.16**	**No**	2.02	
Population									
Neurotypical	13	47	0.21	0.02	0.39	.03	Yes	343.74	86.60
Neurodivergent	3	11	0.39	-0.25	1.03	.23	No	73.30	86.40
Between						**.59**	**No**	0.28	

Skill Classification 1									
Fine	12	43	0.11	-0.11	0.33	.33	No	313.32	86.60
Gross	4	15	0.44	0.13	0.75	.01	Yes	77.47	81.90
Between						**.09**	**No**	2.89	
Skill Classification 2									
Continuous	2	10	0.26	0.05	0.47	.02	Yes	31.04	71.00
Discrete	14	48	0.23	0.00	0.45	.046	Yes	375.65	87.50
Between						**.83**	**No**	0.05	

Note. N = number of studies, K = number of effect sizes, SMD = standardized mean difference, 95% CI = lower and upper confidence intervals, P = significance value, Q = measure of variation around average, I^2 *= measure of proportion of observed variance (%).*

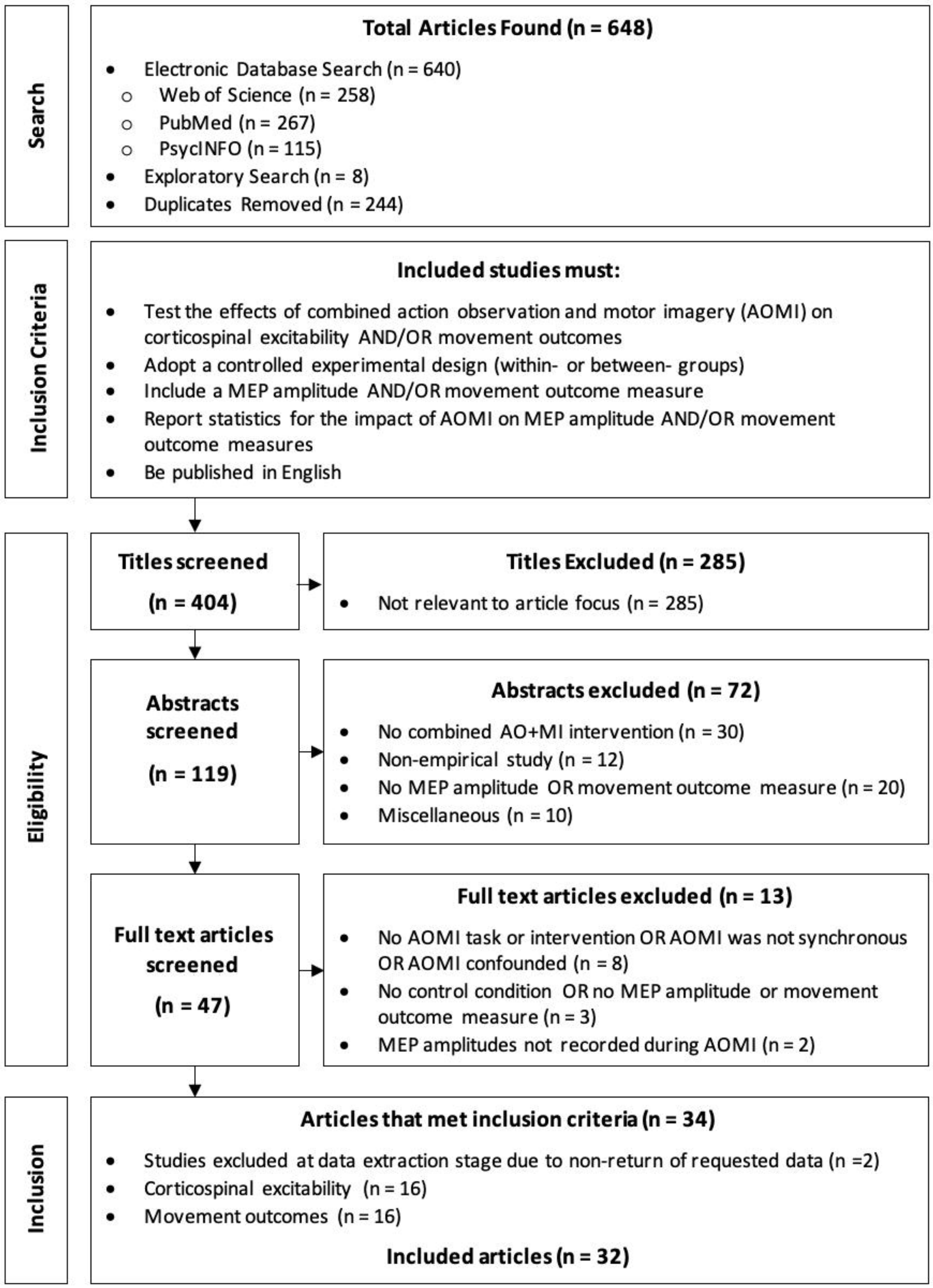
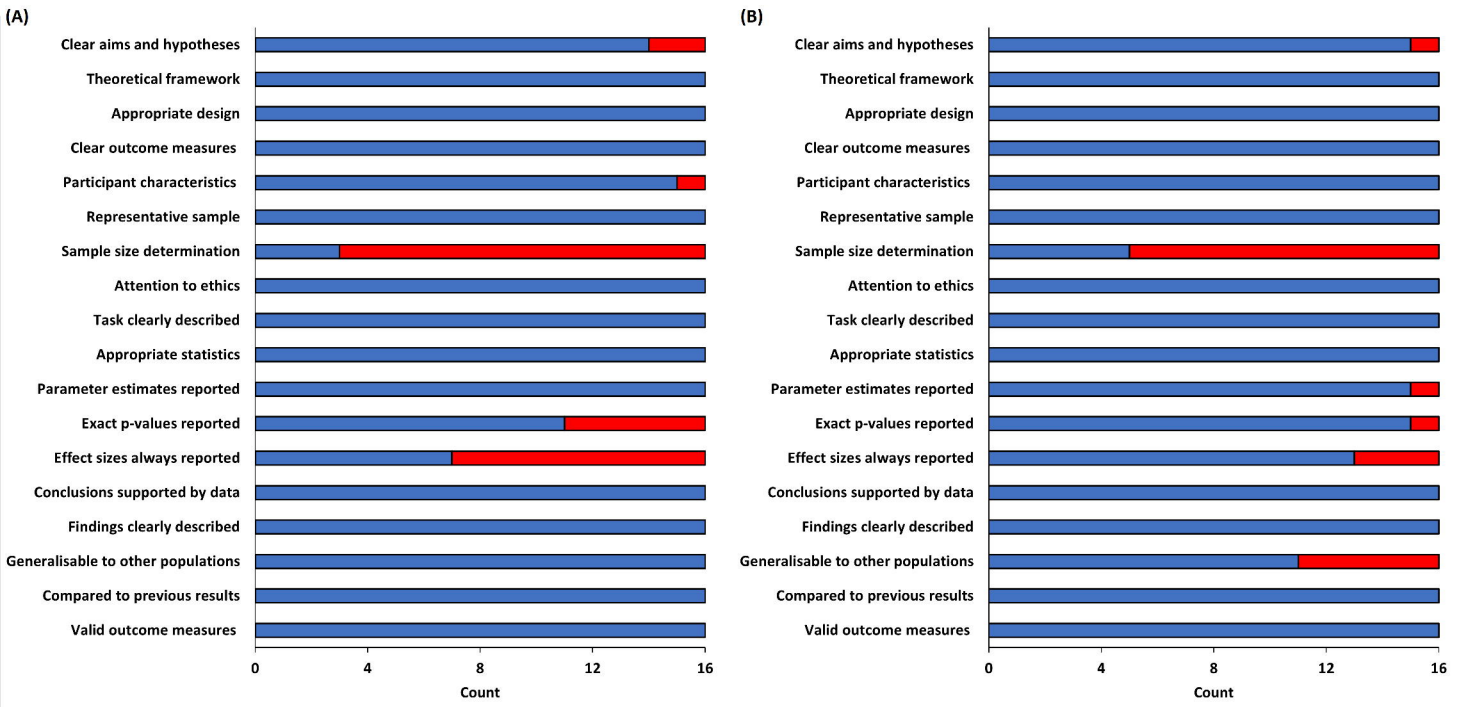
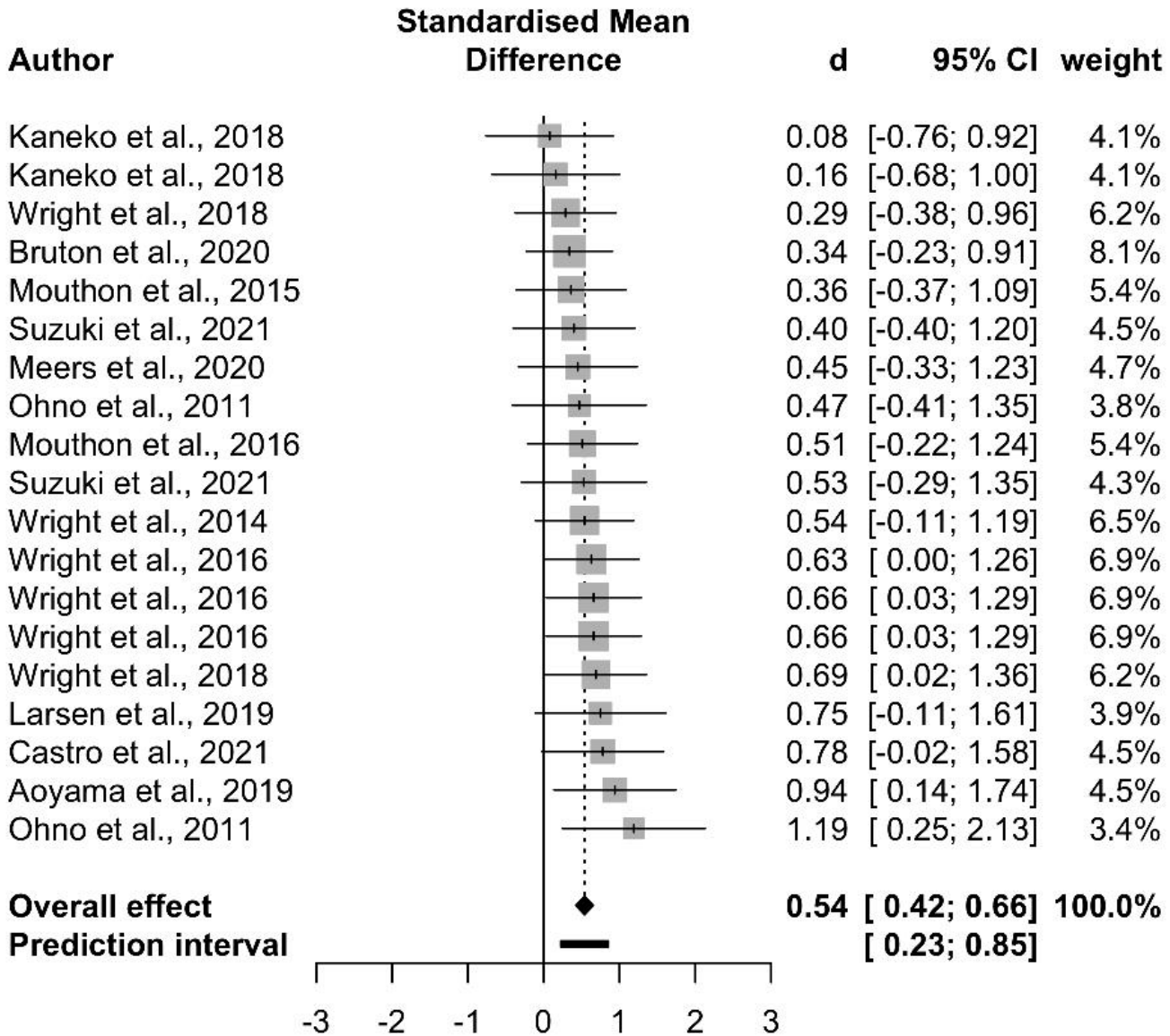
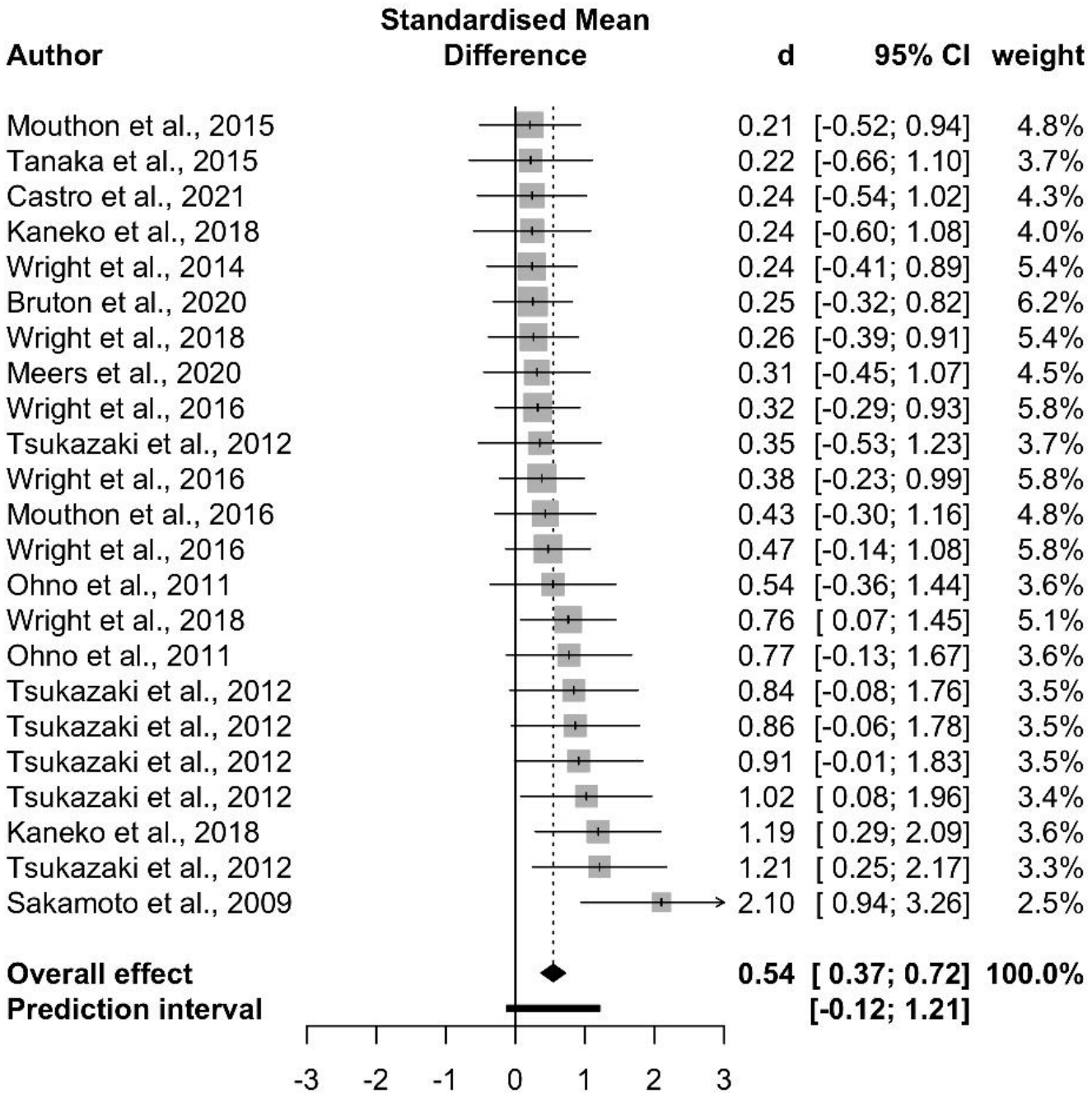
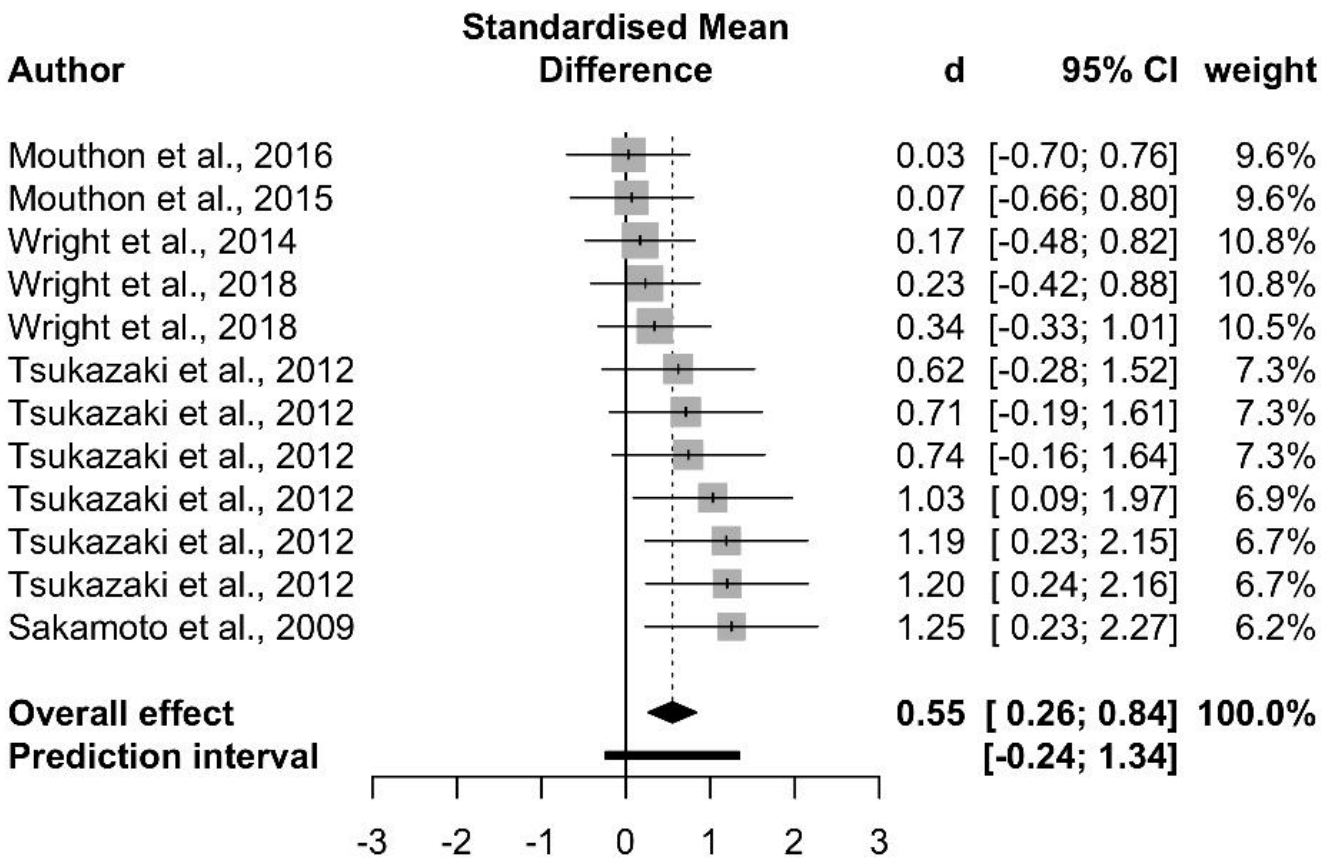


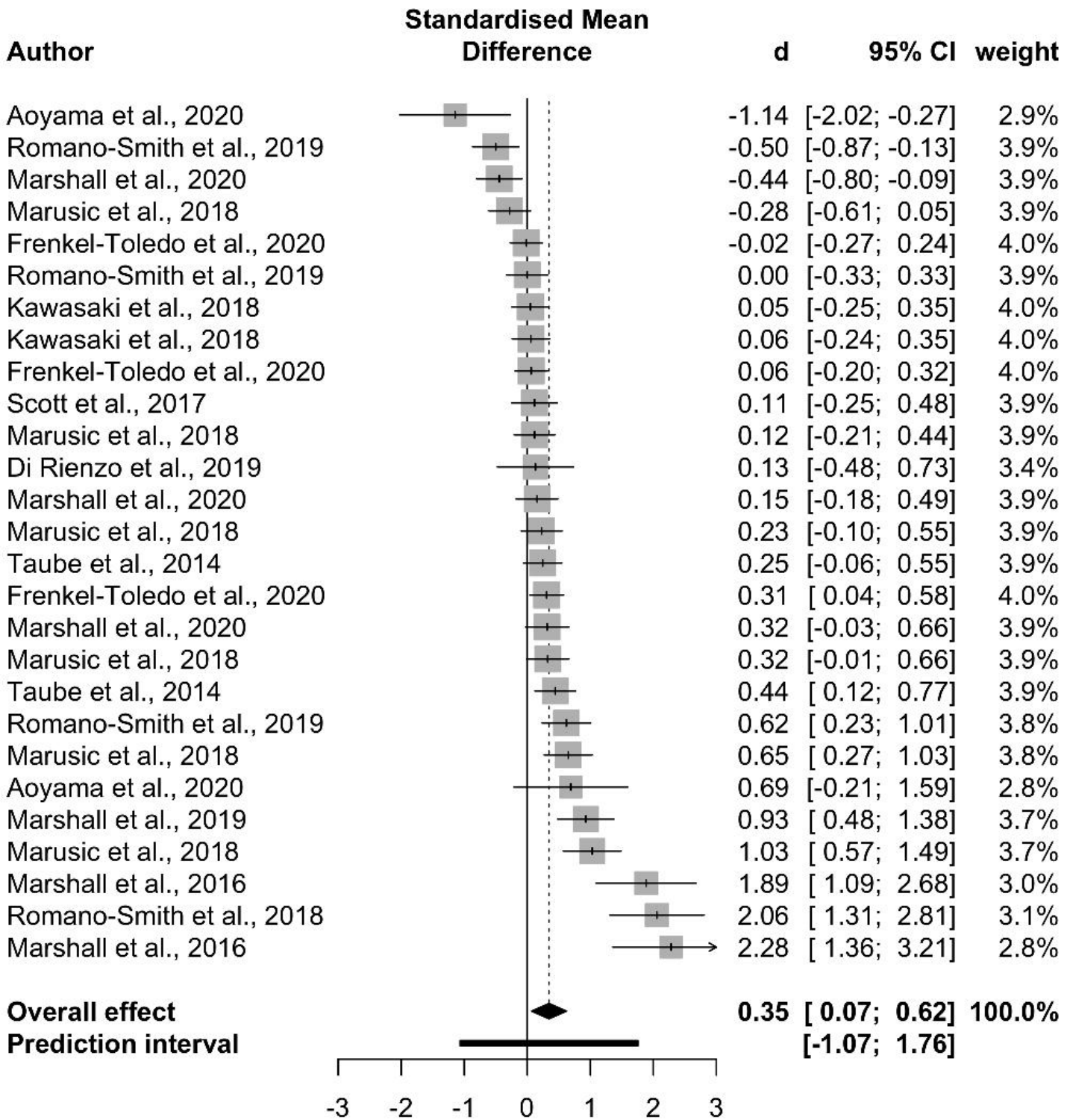
Figure 1. Flow chart of the literature search

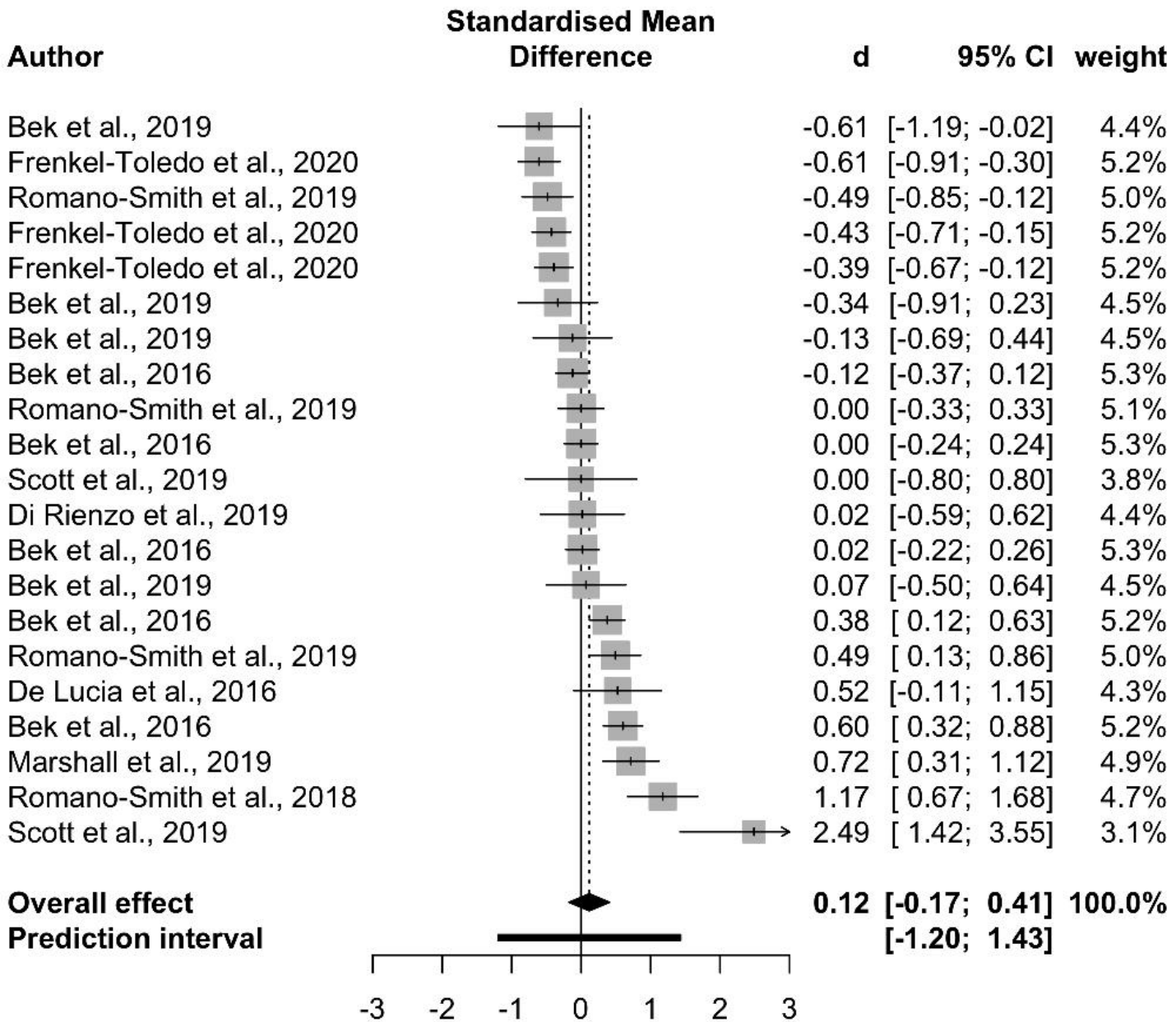

(A)
Clear aims and hypotheses
Theoretical framework
Appropriate design
Clear outcome measures
Participant characteristics
Representative sample
Sample size determination
Attention to ethics
Task clearly described
Appropriate statistics
Parameter estimates reported
Exact p-values reported
Effect sizes always reported
Conclusions supported by data
Findings clearly described
Generalisable to other populations
Compared to previous results
Valid outcome measures
0
4
8
12
16
Count
(B)
Clear aims and hypotheses
Theoretical framework
Appropriate design
Clear outcome measures
Participant characteristics
Representative sample
Sample size determination
Attention to ethics
Task clearly described
Appropriate statistics
Parameter estimates reported
Exact p-values reported
Effect sizes always reported
Conclusions supported by data
Findings clearly described
Generalisable to other populations
Compared to previous results
Valid outcome measures
0
4
8
12
16
Count

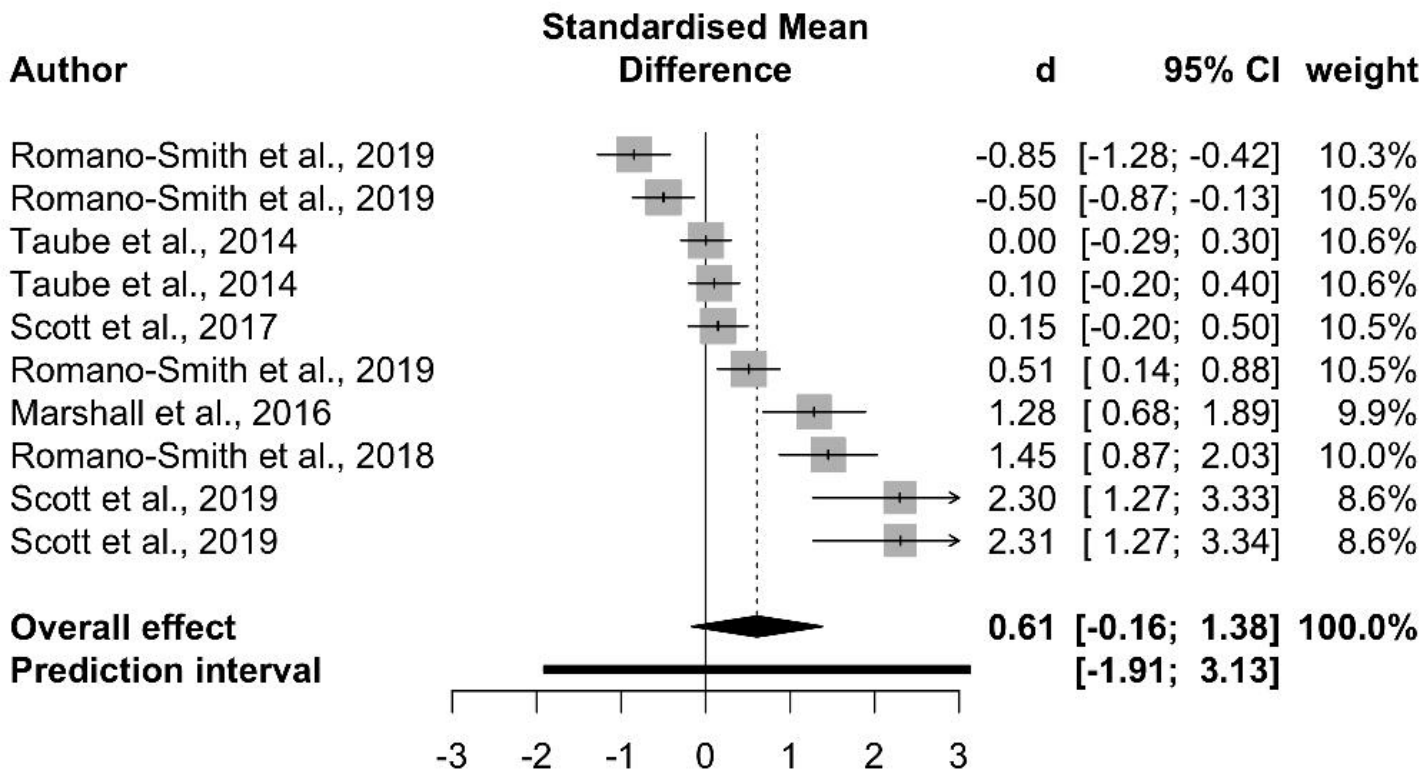

Author
Standardised Mean Difference
d
95% CI
weight
Kaneko et al., 2018 0.08 [-0.76; 0.92] 4.1%
Kaneko et al., 2018 0.16 [-0.68; 1.00] 4.1%
Wright et al., 2018 0.29 [-0.38; 0.96] 6.2%
Bruton et al., 2020 0.34 [-0.23; 0.91] 8.1%
Mouthon et al., 2015 0.36 [-0.37; 1.09] 5.4%
Suzuki et al., 2021 0.40 [-0.40; 1.20] 4.5%
Meers et al., 2020 0.45 [-0.33; 1.23] 4.7%
Ohno et al., 2011 0.47 [-0.41; 1.35] 3.8%
Mouthon et al., 2016 0.51 [-0.22; 1.24] 5.4%
Suzuki et al., 2021 0.53 [-0.29; 1.35] 4.3%
Wright et al., 2014 0.54 [-0.11; 1.19] 6.5%
Wright et al., 2016 0.63 [0.00; 1.26] 6.9%
Wright et al., 2016 0.66 [0.03; 1.29] 6.9%
Wright et al., 2016 0.66 [0.03; 1.29] 6.9%
Wright et al., 2018 0.69 [0.02; 1.36] 6.2%
Larsen et al., 2019 0.75 [-0.11; 1.61] 3.9%
Castro et al., 2021 0.78 [-0.02; 1.58] 4.5%
Aoyama et al., 2019 0.94 [0.14; 1.74] 4.5%
Ohno et al., 2011 1.19 [0.25; 2.13] 3.4%
Overall effect 0.54 [0.42; 0.66] 100.0%
Prediction interval [0.23; 0.85]
-3
-2
-1
0
1
2
3


Standardised Mean Difference
Author
d
95% CI
weight
Mouthon et al., 2015 0.21 [-0.52; 0.94] 4.8%
Tanaka et al., 2015 0.22 [-0.66; 1.10] 3.7%
Castro et al., 2021 0.24 [-0.54; 1.02] 4.3%
Kaneko et al., 2018 0.24 [-0.60; 1.08] 4.0%
Wright et al., 2014 0.24 [-0.41; 0.89] 5.4%
Bruton et al., 2020 0.25 [-0.32; 0.82] 6.2%
Wright et al., 2018 0.26 [-0.39; 0.91] 5.4%
Meers et al., 2020 0.31 [-0.45; 1.07] 4.5%
Wright et al., 2016 0.32 [-0.29; 0.93] 5.8%
Tsukazaki et al., 2012 0.35 [-0.53; 1.23] 3.7%
Wright et al., 2016 0.38 [-0.23; 0.99] 5.8%
Mouthon et al., 2016 0.43 [-0.30; 1.16] 4.8%
Wright et al., 2016 0.47 [-0.14; 1.08] 5.8%
Ohno et al., 2011 0.54 [-0.36; 1.44] 3.6%
Wright et al., 2018 0.76 [0.07; 1.45] 5.1%
Ohno et al., 2011 0.77 [-0.13; 1.67] 3.6%
Tsukazaki et al., 2012 0.84 [-0.08; 1.76] 3.5%
Tsukazaki et al., 2012 0.86 [-0.06; 1.78] 3.5%
Tsukazaki et al., 2012 0.91 [-0.01; 1.83] 3.5%
Tsukazaki et al., 2012 1.02 [0.08; 1.96] 3.4%
Kaneko et al., 2018 1.19 [0.29; 2.09] 3.6%
Tsukazaki et al., 2012 1.21 [0.25; 2.17] 3.3%
Sakamoto et al., 2009 2.10 [0.94; 3.26] 2.5%
Overall effect 0.54 [0.37; 0.72] 100.0%
Prediction interval [-0.12; 1.21]
-3 -2 -1 0 1 2 3

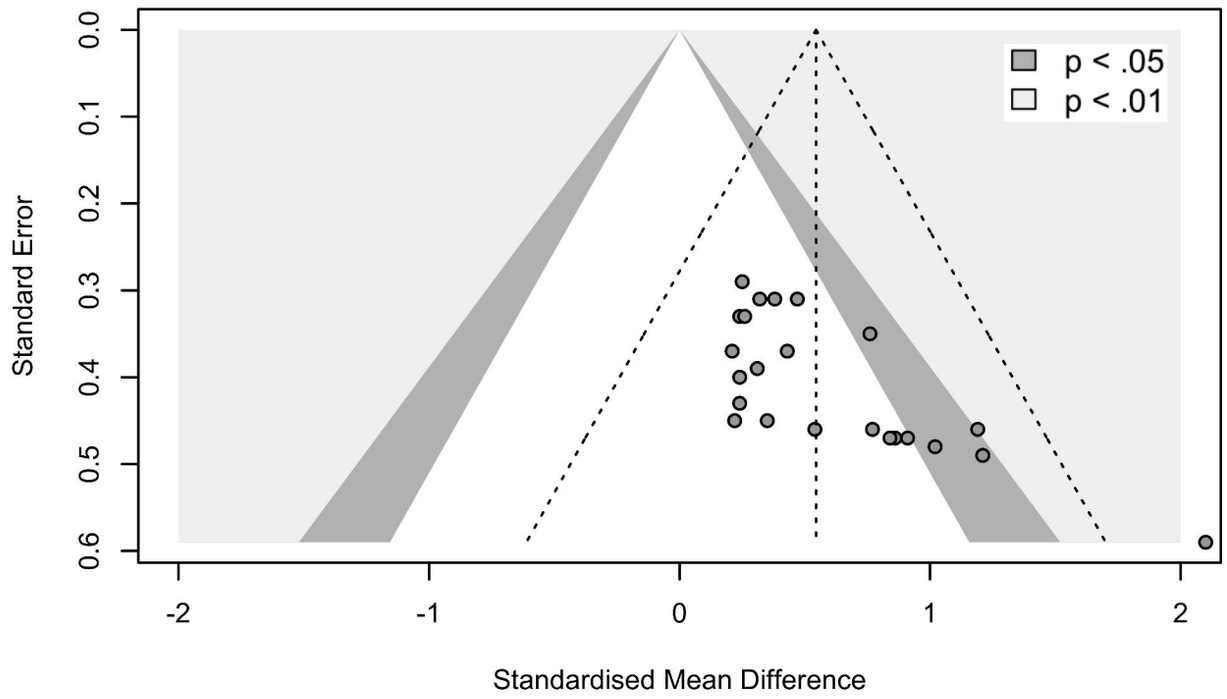

Standardised Mean Difference
Author
d
95% CI
weight
Mouthon et al., 2016 0.03 [-0.70; 0.76] 9.6%
Mouthon et al., 2015 0.07 [-0.66; 0.80] 9.6%
Wright et al., 2014 0.17 [-0.48; 0.82] 10.8%
Wright et al., 2018 0.23 [-0.42; 0.88] 10.8%
Wright et al., 2018 0.34 [-0.33; 1.01] 10.5%
Tsukazaki et al., 2012 0.62 [-0.28; 1.52] 7.3%
Tsukazaki et al., 2012 0.71 [-0.19; 1.61] 7.3%
Tsukazaki et al., 2012 0.74 [-0.16; 1.64] 7.3%
Tsukazaki et al., 2012 1.03 [0.09; 1.97] 6.9%
Tsukazaki et al., 2012 1.19 [0.23; 2.15] 6.7%
Tsukazaki et al., 2012 1.20 [0.24; 2.16] 6.7%
Sakamoto et al., 2009 1.25 [0.23; 2.27] 6.2%
Overall effect 0.55 [0.26; 0.84] 100.0%
Prediction interval [-0.24; 1.34]
-3
-2
-1
0
1
2
3

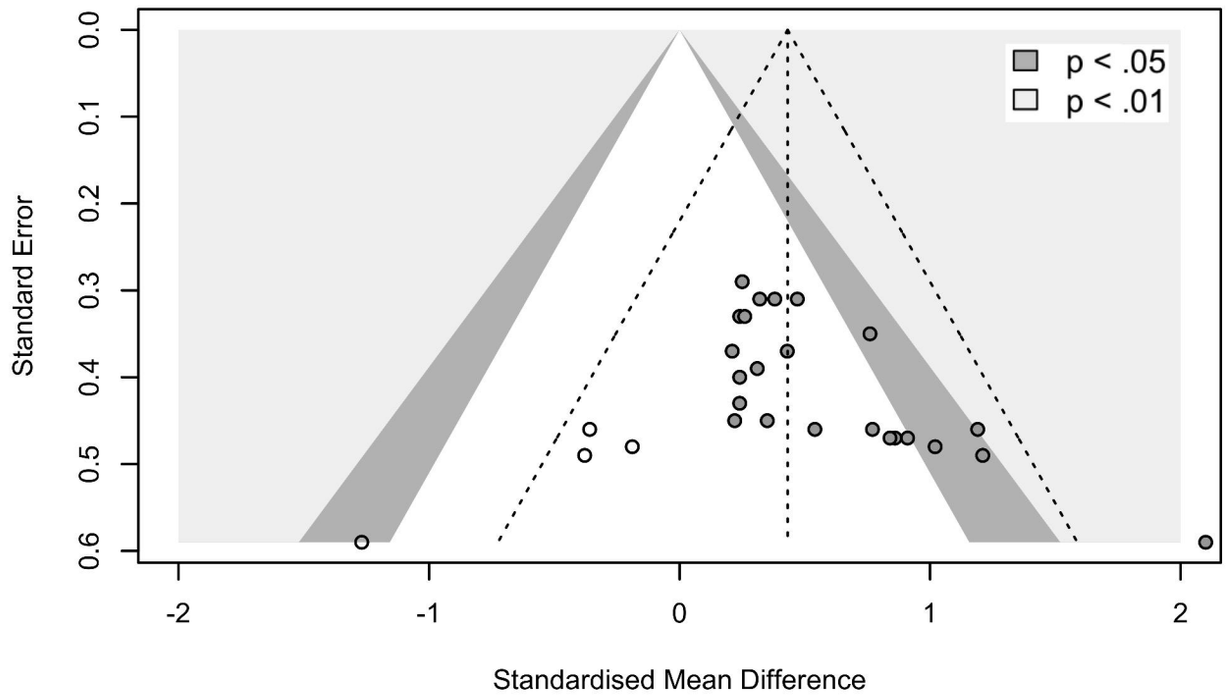

Author
Standardised Mean Difference
d
95% CI
weight
Aoyama et al., 2020 -1.14 [-2.02; -0.27] 2.9%
Romano-Smith et al., 2019 -0.50 [-0.87; -0.13] 3.9%
Marshall et al., 2020 -0.44 [-0.80; -0.09] 3.9%
Marusic et al., 2018 -0.28 [-0.61; 0.05] 3.9%
Frenkel-Toledo et al., 2020 -0.02 [-0.27; 0.24] 4.0%
Romano-Smith et al., 2019 0.00 [-0.33; 0.33] 3.9%
Kawasaki et al., 2018 0.05 [-0.25; 0.35] 4.0%
Kawasaki et al., 2018 0.06 [-0.24; 0.35] 4.0%
Frenkel-Toledo et al., 2020 0.06 [-0.20; 0.32] 4.0%
Scott et al., 2017 0.11 [-0.25; 0.48] 3.9%
Marusic et al., 2018 0.12 [-0.21; 0.44] 3.9%
Di Rienzo et al., 2019 0.13 [-0.48; 0.73] 3.4%
Marshall et al., 2020 0.15 [-0.18; 0.49] 3.9%
Marusic et al., 2018 0.23 [-0.10; 0.55] 3.9%
Taube et al., 2014 0.25 [-0.06; 0.55] 3.9%
Frenkel-Toledo et al., 2020 0.31 [0.04; 0.58] 4.0%
Marshall et al., 2020 0.32 [-0.03; 0.66] 3.9%
Marusic et al., 2018 0.32 [-0.01; 0.66] 3.9%
Taube et al., 2014 0.44 [0.12; 0.77] 3.9%
Romano-Smith et al., 2019 0.62 [0.23; 1.01] 3.8%
Marusic et al., 2018 0.65 [0.27; 1.03] 3.8%
Aoyama et al., 2020 0.69 [-0.21; 1.59] 2.8%
Marshall et al., 2019 0.93 [0.48; 1.38] 3.7%
Marusic et al., 2018 1.03 [0.57; 1.49] 3.7%
Marshall et al., 2016 1.89 [1.09; 2.68] 3.0%
Romano-Smith et al., 2018 2.06 [1.31; 2.81] 3.1%
Marshall et al., 2016 2.28 [1.36; 3.21] 2.8%
Overall effect 0.35 [0.07; 0.62] 100.0%
Prediction interval [-1.07; 1.76]
-3 -2 -1 0 1 2 3


Standardised Mean Difference
Author
d
95% CI
weight
Bek et al., 2019 -0.61 [-1.19; -0.02] 4.4%
Frenkel-Toledo et al., 2020 -0.61 [-0.91; -0.30] 5.2%
Romano-Smith et al., 2019 -0.49 [-0.85; -0.12] 5.0%
Frenkel-Toledo et al., 2020 -0.43 [-0.71; -0.15] 5.2%
Frenkel-Toledo et al., 2020 -0.39 [-0.67; -0.12] 5.2%
Bek et al., 2019 -0.34 [-0.91; 0.23] 4.5%
Bek et al., 2019 -0.13 [-0.69; 0.44] 4.5%
Bek et al., 2016 -0.12 [-0.37; 0.12] 5.3%
Romano-Smith et al., 2019 0.00 [-0.33; 0.33] 5.1%
Bek et al., 2016 0.00 [-0.24; 0.24] 5.3%
Scott et al., 2019 0.00 [-0.80; 0.80] 3.8%
Di Rienzo et al., 2019 0.02 [-0.59; 0.62] 4.4%
Bek et al., 2016 0.02 [-0.22; 0.26] 5.3%
Bek et al., 2019 0.07 [-0.50; 0.64] 4.5%
Bek et al., 2016 0.38 [0.12; 0.63] 5.2%
Romano-Smith et al., 2019 0.49 [0.13; 0.86] 5.0%
De Lucia et al., 2016 0.52 [-0.11; 1.15] 4.3%
Bek et al., 2016 0.60 [0.32; 0.88] 5.2%
Marshall et al., 2019 0.72 [0.31; 1.12] 4.9%
Romano-Smith et al., 2018 1.17 [0.67; 1.68] 4.7%
Scott et al., 2019 2.49 [1.42; 3.55] 3.1%
Overall effect 0.12 [-0.17; 0.41] 100.0%
Prediction interval [-1.20; 1.43]
-3 -2 -1 0 1 2 3

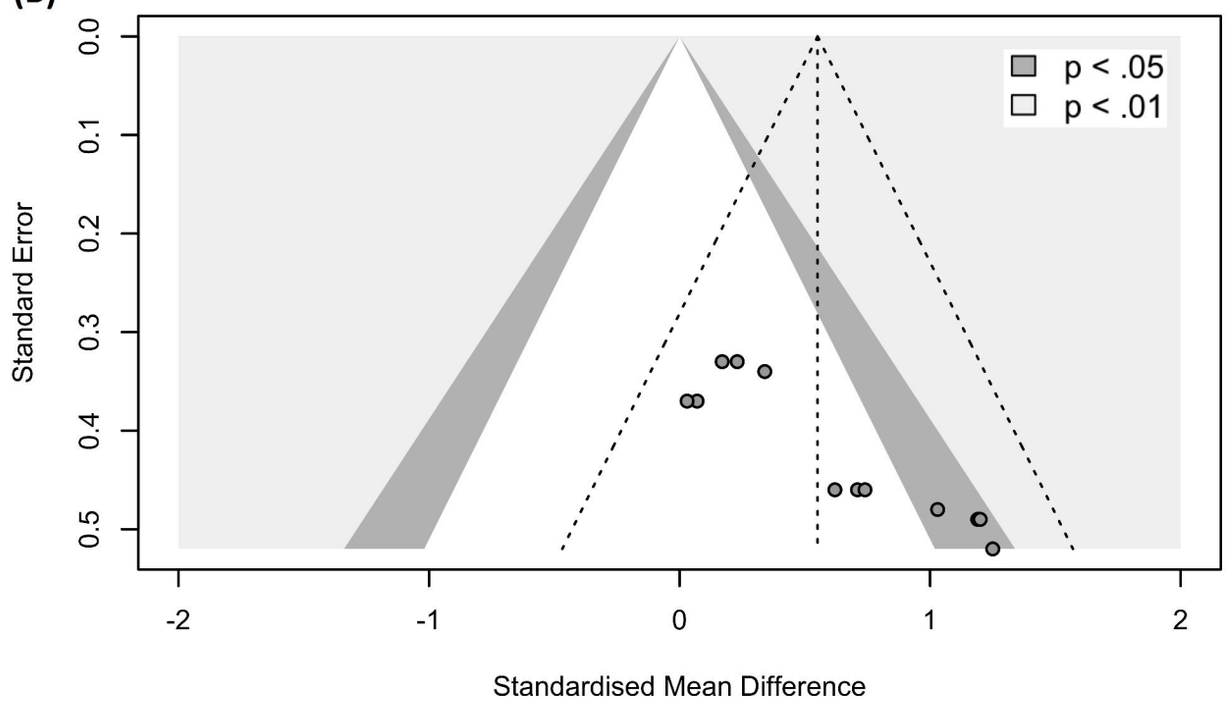

Standardised Mean Difference
Author
d
95% CI
weight
Romano-Smith et al., 2019 -0.85 [-1.28; -0.42] 10.3%
Romano-Smith et al., 2019 -0.50 [-0.87; -0.13] 10.5%
Taube et al., 2014 0.00 [-0.29; 0.30] 10.6%
Taube et al., 2014 0.10 [-0.20; 0.40] 10.6%
Scott et al., 2017 0.15 [-0.20; 0.50] 10.5%
Romano-Smith et al., 2019 0.51 [0.14; 0.88] 10.5%
Marshall et al., 2016 1.28 [0.68; 1.89] 9.9%
Romano-Smith et al., 2018 1.45 [0.87; 2.03] 10.0%
Scott et al., 2019 2.30 [1.27; 3.33] 8.6%
Scott et al., 2019 2.31 [1.27; 3.34] 8.6%
Overall effect 0.61 [-0.16; 1.38] 100.0%
Prediction interval [-1.91; 3.13]
-3 -2 -1 0 1 2 3

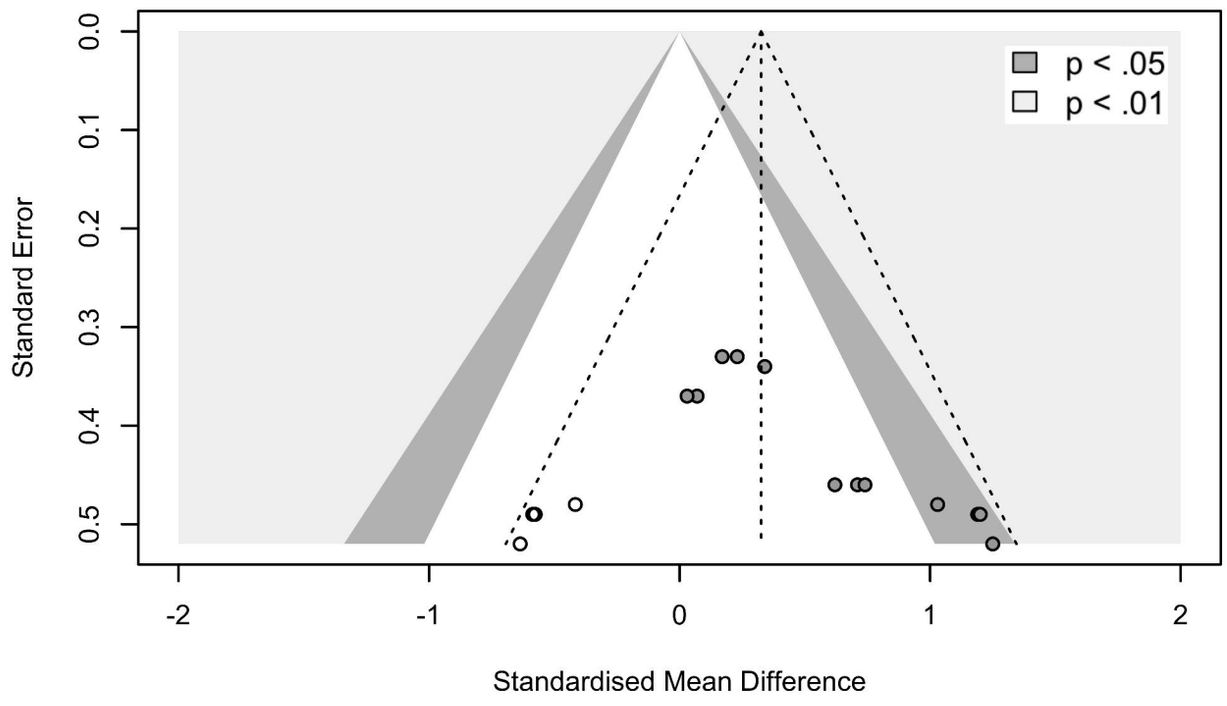
(A)


p < .05
p < .01
Standard Error
0.0
0.1
0.2
0.3
0.4
0.5
0.6
-2
-1
0
1
2
Standardised Mean Difference

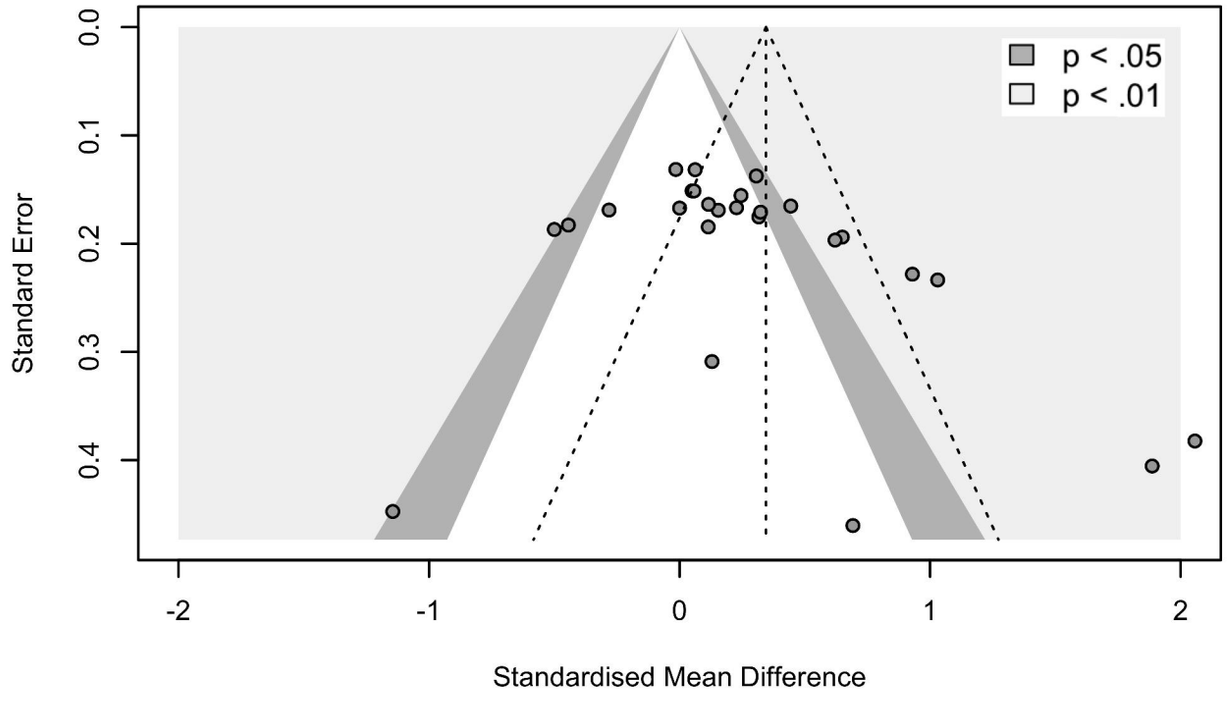

p < .05
p < .01
Standard Error
0.0
0.1
0.2
0.3
0.4
0.5
0.6
-2
-1
0
1
2
Standardised Mean Difference

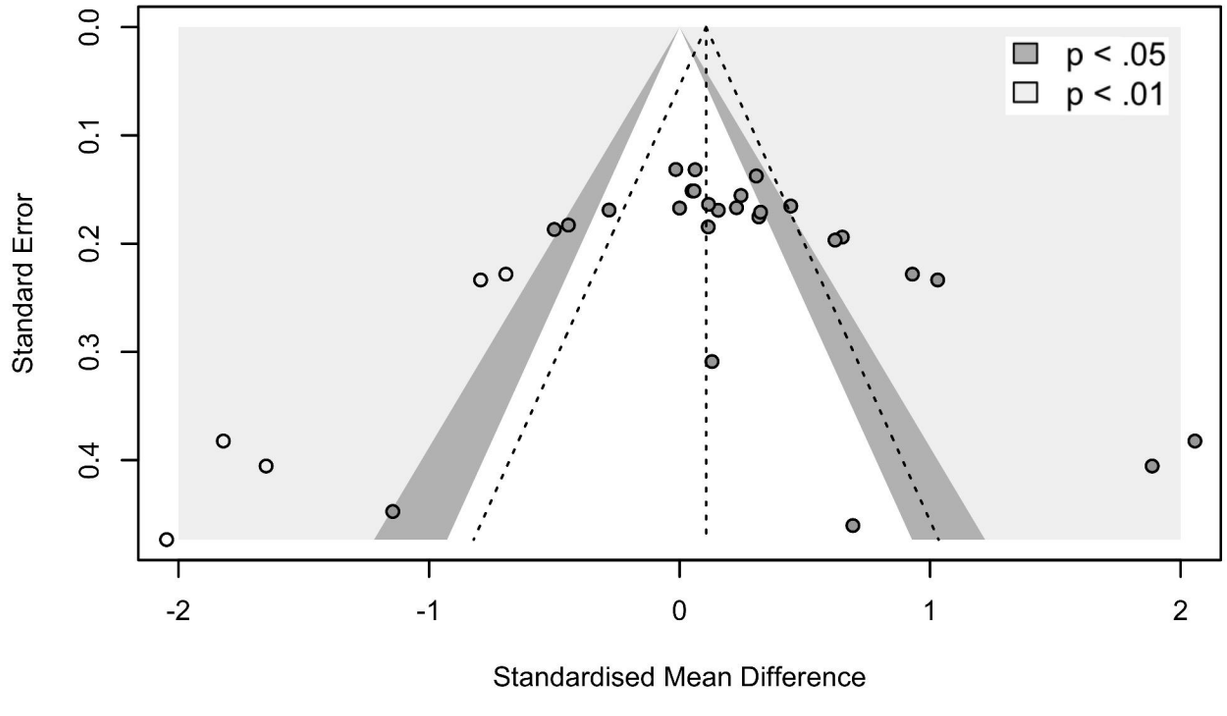
(B)


p < .05
p < .01
Standard Error
0.0
0.1
0.2
0.3
0.4
0.5
-2
-1
0
1
2
Standardised Mean Difference


p < .05
p < .01
Standard Error
0.0
0.1
0.2
0.3
0.4
0.5
-2
-1
0
1
2
Standardised Mean Difference

(C)


p < .05
p < .01
Standard Error
0.0
0.1
0.2
0.3
0.4
-2
-1
0
1
2
Standardised Mean Difference


p < .05
p < .01
Standard Error
0.0
0.1
0.2
0.3
0.4
-2
-1
0
1
2
Standardised Mean Difference