Development and Validation of the Combined Action Observation and Motor Imagery Ability Questionnaire

Matthew W. Scott[1,2], Maaike Esselaar[3] Neil Dagnall[4], Andrew Denovan[5], Ben Marshall[3],

Aimee S. Deacon[4], Paul S. Holmes[3] & David J. Wright[4*]

Affiliations

1. School of Kinesiology, University of British Columbia, Vancouver, Canada.
2. Department of Psychology, University of British Columbia, Kelowna, Canada.
3. Department of Sport and Exercise Sciences, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, UK
4. Department of Psychology, Faculty of Health and Education, Manchester Metropolitan University, Manchester, UK
5. School of Psychology, Liverpool John Moores University, Liverpool, UK

*** Corresponding author**

Dr. David J. Wright

Department of Psychology, Manchester Metropolitan University, Manchester, UK.

Email: d.j.wright@mmu.ac.uk

Abstract

Combined use of action observation and motor imagery (AOMI) is an increasingly popular motor simulation intervention, which involves observing movements on video whilst simultaneously imagining the feeling of movement execution. Measuring and reporting participant imagery ability characteristics is essential in motor simulation research, but no measure of AOMI ability currently exists. Accordingly, the AOMI ability questionnaire (AOMI-AQ) was developed to address this gap in the literature. In Study 1, 211 participants completed the AOMI-AQ and the kinesthetic imagery sub-scales of the Movement Imagery Questionnaire-3 and Vividness of Motor Imagery Questionnaire-2. Following exploratory factor analysis, an 8-item AOMI-AQ was found to correlate positively with existing motor imagery measures. In Study 2, 174 participants completed the AOMI-AQ a second time 7-10 days later. Results indicate a good test-retest reliability for the AOMI-AQ. The new AOMI-AQ measure provides a valid and reliable tool for researchers and practitioners wishing to assess AOMI ability.

Keywords: Motor simulation, imagery ability, motor imagery during action observation, scale development

Introduction

Motor imagery (MI) is a perceptual-like experience involving the mental generation, manipulation and maintenance of the visual and kinesthetic properties of a movement (Kosslyn et al., 2010). Motor imagery ability, therefore, refers to how well an individual can create, maintain and control their motor imagery (Morris et al., 2005). MI is used widely in psychological intervention programmes to improve physical performance (e.g., Robin et al., 2023; for a meta-analysis see Toth et al., 2020) and enhance psychological processes related to performance (e.g., motivation; Simonsmeier et al., 2021). However, within the last decade, imagery research has turned to alternative methods of simulation-based training. One method which has gained interest is the combined and simultaneous use of action observation and motor imagery (AOMI). AOMI involves watching a video or live display of an action, while generating, maintaining, and transforming a time-synched kinesthetic representation of the same action (Scott et al., 2022). In practice, AOMI therefore involves watching movements whilst imagining concurrently the kinesthetic sensations of executing the observed action. Although instructions are typically limited to kinesthetic imagery during AOMI, this does not exclude the spontaneous use of, or requirement for, visual imagery (VI; see Mizuguchi et al., 2016; Scott et al., 2022); however, the visual display during AOMI reduces the requirement for VI to some extent (Wright et al., 2022).

Though some studies almost 20 years ago used an approach resembling AOMI (e.g., Smith & Holmes, 2004), the technique has gained increased attention within the last decade (Eaves et al., 2016a; Vogt et al., 2013). Research into AOMI has focused on exploring its neurophysiological or behavioural effects in relation to independent MI or AO (Eaves et al., 2016a; 2022; O'Shea, 2022; Scott et al., 2021; 2022). Recent meta-analytical findings indicate that AOMI elicits increased activity in areas of the brain related to motor planning and execution and produces superior performance outcomes compared to independent AO, and effects are at least equivalent to MI on both outcome measures (Chye et al., 2022).

There are several theoretical and practical constraints associated with MI that can be resolved by AOMI. For example, as many as 4% of individuals report being unable to generate VI or experience difficulties in doing so (i.e., Aphantasia; Dance et al., 2022). This issue could potentially be addressed through the provision of visual stimuli on video during AOMI. In addition, due to the covert nature of imagery, practitioners have limited control over their client's imagery experience (i.e., visual perspective, viewing angle, movement timing, and image maintenance; Holmes & Calmels, 2008). These problems may also be resolved through AOMI since the practitioner can control the visual perspective, viewing angle, and movement timing information through the video stimuli they present to their client (Wright et al., 2022). Furthermore, content related to the task and environment, which may vary across individuals based on previous motor or visual experience (Malouin et al., 2009; Wright et al., 2015) are controlled in AOMI (Scott et al., 2022). Another advantage of AOMI over MI is the opportunity for controlling the agent of the simulated action (i.e., via self- or other-modelling). This is important because some individuals may have a natural preference for imagining themselves or others (Holmes & Calmels, 2008).

A challenge that remains for researchers and practitioners using AOMI is that no measure of AOMI ability currently exists, despite growing interest in how AOMI ability may influence performance (e.g., McNeill et al., 2020; Robin & Blandin, 2021). As such, researchers appear to have assumed implicitly that all individuals can engage easily in AOMI. Indeed, in Vogt et al.'s (2013) seminal paper, the authors asserted that AOMI "does not take particular skill" (p. 1). Though this may be true for many individuals, given known differences in VI ability (Dance et al., 2022), reduced imagery ability across the lifespan (Gulyas et al., 2022; Spruijt et al., 2015), and in clinical populations (e.g., de Vries et al., 2013; Emerson et al., 2018; la Touche et al., 2015; Scott et al., 2021), this is not necessarily the case.

In the absence of an appropriate measure of AOMI ability, much of the research to date has failed to include any assessment of participants' imagery abilities (e.g., Kawasaki et al., 2018; Marshall et al., 2020; Rungsirisilp & Wongsawat, 2022; Taube et al., 2014; 2015). This is problematic as recent guidelines recommend strongly that a measure of imagery ability should be included when reporting AOMI studies (Moreno-Verdú et al., 2022). Alternatively, some investigators have used existing MI ability questionnaires as a proxy measure of AOMI ability (e.g., Emerson et al., 2022; McNeill et al., 2021; Romano-Smith et al., 2022; Scott et al., 2017; Wright et al., 2018). Although this approach seems intuitive in the absence of an existing measure of AOMI ability, this may not be entirely appropriate as AOMI and MI may rely on relatively different neural substrates and utilize different cognitive processes (O'Shea, 2022; Scott et al., 2022). For instance, AOMI results in greater activations bilaterally across the primary motor cortex and cerebellum and stronger activity in the precuneus than independent MI (Taube et al., 2015). Furthermore, the greater recruitment of the rostral pre-frontal cortex during AOMI may indicate different cognitive requirements for its use (Eaves et al., 2016b; Emerson et al., 2022). Conceptually, having the ability to generate and manipulate visual and kinesthetic imagery may not necessarily translate to being able to generate and maintain a kinesthetic imagery (KI) representation in synchrony with an external video or live demonstration. Indeed, it has been proposed that in contrast to typical deliveries of MI, using AOMI requires attentional switching between internal (KI) and external (AO) components (Eaves et al., 2016b; Emerson et al., 2022). Considering these factors, together with the current research interest in AOMI processes, there is a clear need to develop an appropriate measure of AOMI ability.

Though no measures of AOMI ability presently exist, the creation of a new instrument can be informed by existing indices of MI ability. It has been proposed that imagery ability consists of different processes, such as the ability to generate, manipulate or maintain imagined content (Eaves & Cummings, 2018; Kraeutner et al., 2020). Several measures of imagery ability have been developed, including implicit (i.e., hand laterality judgement task) and explicit (i.e., mental chronometry or self-report) measures, which may probe these different imagery processes (Kraeutner et al., 2020). However, AOMI versions of measures like the hand laterality judgement task or mental chronometry would not be appropriate, as video stimuli would inherently depict the mental rotation and movement timing information being measured in these respective tasks. Therefore, a self-report measure for AOMI ability is most appropriate and has the added benefit of being easy to administer.

One of the most widely used measures of MI ability is the Vividness of Movement Imagery Questionnaire-2 (VMIQ-2; Roberts et al., 2008). The VMIQ-2 instructs individuals to imagine 12 actions from first- and third-person visual perspectives and a separate kinesthetic modality, all of which show excellent internal reliability (αs > 0.93; Roberts et al., 2008). For each action, participants are required to rate the vividness of their imagined movement on a 5-point Likert scale. As the questionnaire imposes no requirement to execute the movements physically prior to imagining them, the movements imagined could conceivably be either within or beyond the individual's own motor repertoire. Williams et al. (2012) previously highlighted limitations in the delivery of this measure due to the open interpretation of some of its items. One item, for example, requires the individual to imagine kicking a ball in the air. This could, conceivably, introduce variation in responses based on the type of ball or kick imagined, which is potentially biased by an individual's previous motor or visual experiences. Other items instruct people to imagine actions that they may never have performed, such as swinging on a rope or jumping off a high wall. An AOMI version of this questionnaire, with the requirement for participants to rate how easily they can imagine the feeling of an action they may have never performed, would clearly be problematic.

An alternative measure adopted within MI research is the Movement Imagery Questionnaire-3 (MIQ-3; Williams et al., 2012). Like the VMIQ-2, the MIQ-3 measures KI and VI from both first- and third-person perspectives, by requiring participants to rate the ease with which they can generate the imagined movement. Furthermore, the MIQ-3 shows good composite reliability across first- and third-person VI and for KI (all values > 0.79; Williams et al., 2012). However, in contrast to the VMIQ-2, the MIQ-3 instructs the individual to perform the movement immediately before imagining it. This ensures that participants acquire motor and first-person visual experience of each action immediately prior to imagining it, overcoming the limitation of the VMIQ-2 regarding previous experiences. Accordingly, the MIQ-3 provides more control over potential confounds that may occur within the imagery process, allowing a more direct measure of imagery generation ability, and providing a suitable template from which to develop an AOMI ability questionnaire.

The aim of the first study was to construct and validate a self-report measure of AOMI ability – the combined action observation and motor imagery ability questionnaire (AOMI-AQ). To achieve this, we developed the AOMI-AQ and conducted an exploratory factor analysis examining the underlying structure of the measure, along with validation of the AOMI-AQ against established measures of KI ability; the VMIQ-2 and MIQ-3 (Study 1). The second study then established the test-retest reliability of the AOMI-

AQ measure and investigated response variation depending on supervised or unsupervised completion of the measure (Study 2).

Study 1

Methods

Participants

Two-hundred and eleven participants (F = 121, M = 89, non-binary = 1) aged between 18 to 57 years (*M* = 29, *SD* = 9.03), with normal or corrected-to-normal vision, participated in Study 1. The relatively wide age range in the sample was appropriate given the ongoing use of AOMI across younger and older age groups (e.g., Emerson et al., 2022; Kawasaki et al., 2018; Mouthon et al., 2020; see Chye et al., 2022 for meta-analysis). Of this sample, 188 participants self-reported being right-handed, with the remainder identifying as left-hand dominant. All participants self-reported being able to perform the actions required to complete the MIQ-3 and AOMI-AQ. The current sample size was sufficient for exploratory factor analysis based on previous guidelines by Comrey and Lee (1992). Furthermore, the sample size exceeded that used in other studies that have adapted the MIQ-3 for use with children (Martini et al., 2016) or for use in different languages (e.g., Dilek et al., 2020; Trapero-Asenjo et al., 2021). Ethical approval for both studies was granted through the Manchester Metropolitan University Faculty of Health and Education Research Ethics and Governance Committee (Ethical Approval Identification Number: 35170) and participants provided written informed consent prior to participating.

Materials

Development of the combined action observation and motor imagery ability questionnaire (AOMI-AQ). A Qualtrics and Microsoft PowerPoint version of the AOMI-AQ can be accessed and downloaded via https://osf.io/vbqjw/?view_only=3382b7e43a794ed78ea0c17a17eebe1f . A PsychoPy version can be accessed by emailing the authors. The AOMI-AQ was developed by adapting the MIQ-3 into a video-based questionnaire depicting similar actions, recorded from both first- and third-person visual perspectives. The MIQ-3 comprises four movements, which individuals perform and subsequently imagine: a right knee raise, a crouch to jump, a horizontal arm adduction (non-dominant limb), and a waist bend toe touch movement. The same movements were used for the AOMI-AQ although the arm action was modified to include an initial vertical abduction, followed by horizontal adduction. This decision was taken to provide additional visual movement information with which participants could synchronize their imagery during AOMI. A library of videos was created with both a male (Caucasian, 28 years old) and female (Caucasian, 27 years old) model. Different sex models were included in agreement with AOMI guidelines (Wright et al., 2022), to allow participants to perform AOMI with the video model who matched their chosen gender identity most closely. Each movement was timed to a metronome and filmed from both first- and third-person visual perspectives and showing the models performing as if right- or left-hand dominant, as required for the arm raise action.

First-person perspective videos were recorded from a head mounted camera to capture the movement from the viewpoint that most accurately represented the view participants would see if performing the movement themselves. Third-person perspective videos were filmed depicting the model from a 45-degree angle. This angle was chosen in favor of the sagittal or frontal plane due to opportunities to highlight lateral, anterior and posterior components of the movements. For example, a frontal view of a knee raise would accurately depict lateral sway when shifting onto one leg but would not accurately show movement range and depth in the anterior/posterior plane. Movements were recorded from both first- and third-person visual perspectives as each perspective may have benefits for AOMI. For example, a first-person perspective can give the illusion of agency (Scott et al., 2022), whereas AOMI from third-person perspectives could provide valuable task relevant visual cue information regarding movement form, which is not present during the first-person perspective (Hardy & Callow, 1999; Holmes & Collins, 2001; Scott et al., 2022). As such, incorporating both perspectives likely captured perspective dependent differences across the tasks.

The AOMI-AQ required participants to report how well they were able to generate KI whilst observing a video of each action. In contrast to the MIQ-3, however, no VI rating scale was included in the AOMI-AQ as VI is not typically instructed during AOMI (Scott et al., 2022; Wright et al., 2022). An additional subscale was included requiring participants to rate how well they were able to maintain the imagined feeling throughout the video. As the questionnaire movements ranged from slower, smoother movements (e.g., a knee raise) to more explosive movements (e.g., a crouch-to-jump), the generation of KI throughout movements would presumably fluctuate due to different muscle engagement, force requirements, and movement durations during these movements. For instance, in the crouch-to-jump movement KI should theoretically be strongest during crouch and extension components, and on landing when kinesthetic feedback should be most salient during actual performance, but with limited KI during the flight phase. Maintenance of KI for this jump movement could therefore be expected to differ from that of the other three movements. The inclusion of a maintenance subscale allowed these potential differences to be quantified.

Completion of the combined action observation and motor imagery ability questionnaire (AOMI-AQ). Participants first selected whether the male or female model most closely matched their gender identity, to ensure that the appropriate videos played where relevant during completion of the questionnaire. As illustrated in Figure 1, for each item participants first saw a movement demonstration from both the first- and third-person perspective. The order of these perspectives alternated based on the perspective from which AOMI would be performed – if AOMI was in the third-person they would see a third-, then first-person perspective demonstration before physical performance and AOMI of the movement, and vice versa. During the MIQ-3 participants would typically be provided with a written description of the to-be performed and imagined action. In contrast the AOMI-AQ provides video demonstrations of these actions to reduce the likelihood of participants misinterpreting the movements, and furthermore, this may improve the efficacy for unsupervised delivery of the questionnaire. As illustrated in Figure 1, participants saw both perspectives in a counterbalanced order depending on the perspective of the item, this ensured that: 1) participants knew the content they would be required to imagine and exactly how the movement should be performed, and 2) participants were not overly exposed to one visual perspective throughout the questionnaire. Although the third-person perspective would presumably provide more salient information regarding the movement form (Scott et al., 2021), only showing this perspective for demonstrations may have led to a response bias for third-person perspective

AOMI items. Furthermore, exposing participants to the exact same video immediately prior to the true AOMI assessment video may provide practice which could influence or inflate responses.

Figure 1. Timeline for first-person (1pp) and third-person perspective (3pp) items. Participants were shown a demonstration of the movement they would perform during AOMI from first-person (1pp) and third-person (3pp) visual perspectives. Perspective order during the demonstrations was dependent on AOMI perspective. For example, if AOMI was to be performed in the first-person perspective the participant saw a first- then third-person person perspective demonstration (top panel). This order was reversed for third-person perspective AOMI items. The participant would then perform the movement physically before engaging in AOMI and then rating their ease of image generation or maintenance.

Participants then saw written instructions prompting them to perform one attempt of the movement they had just seen. After executing the movement, they were instructed to stand in the starting position for the action and to watch a video of that action while simultaneously imagining the feeling of the movement in time with the video and as if they were performing it. Finally, using a 7-point Likert-type scale participants reported the difficulty/ease of their KI generation (1 – "*Very hard to feel*" to 7 – "*Very easy to feel*") or the difficulty/ease of their KI maintenance (1 – "*Very hard to maintain*" to 7 – "*Very easy to maintain*"). Accounting for the two subscales (generate and maintain), two visual perspectives (first- and third-person) and the four movements, this created a 16-item questionnaire. The order the movements were delivered was identical to the MIQ-3; however, video perspectives and subscale questions (generation or maintenance) were delivered in an interleaved fashion.

Movement imagery questionnaire-3 (MIQ-3; Williams et al., 2012). Participants completed the kinesthetic imagery subscale of the MIQ-3. The VI subscales of the MIQ-3 were omitted given that the AOMI-AQ does not measure this modality (due to limited requirements for VI), and only KI being instructed during AOMI. This subscale consists of four actions, which participants are required imagine:

1) a waist bend (toe touch), 2) crouch and jump, 3) an arm movement (non-dominant arm), and 4) knee raise. For each item, participants performed the action once before they imagined the feeling as if they were again performing the movement. Participants then rated how well they were able to imagine the feeling of the movement on a 7-point Likert type scale (1 – "*Very hard to feel*" to 7 – "*Very easy to feel*"). The MIQ-3 has good psychometric properties, internal reliability, and predictive validity (composite reliability ≥0.7 for all subscales; Williams et al., 2012).

Vividness of movement imagery questionnaire-2 (VMIQ-2; Roberts et al., 2008). Like the MIQ-3, participants only completed the KI subscale of the VMIQ-2, which consisted of 12 items. Items included imagining different movements ranging from more isolated tasks (e.g., throwing a stone into water) to more full body movements (e.g., jumping off a high wall). For each item participants were asked to rate how well they were able imagine the feeling of performing the movement. The Likert type scale for the VMIQ-2 ranges from 1 ("*No image at all, you only know you are thinking of the skill*") to 5 ("*Perfectly clear and vivid as normal feel of movement*"). Note that this scoring scale is reversed from the original VMIQ-2. This ensured that higher scores indicate higher imagery ability and established consistency with the scale orientation of the MIQ-3 and AOMI-AQ to reduce the likelihood of misinterpretation. The VMIQ-2 has good psychometric properties, internal reliability, and predictive validity (Roberts et al., 2008).

Procedures

Prior to undertaking the study participants were informed of the study aims and provided written informed consent. Next, participants completed a demographic questionnaire in which they reported their age, gender identity, ethnicity, handedness, and whether a male or female model most closely matched their gender identity. Following this, participants were educated on the concept of imagery with an emphasis on KI. Once participants understood the content they would be required to imagine, they completed the AOMI-AQ, VMIQ-2 and MIQ-3 in the presence and under the supervision of a member of the research team, who supervised the testing session and ensured that participants adhered to the instructions to execute and imagine the movements when required. All questionnaires were delivered through PsychoPy (V2023.1.1, Peirce et al., 2019), displayed from a laptop, and the delivery of these questionnaires was counterbalanced across participants to control order effects.

Depending on responses to the demographic questionnaire, during completion of the AOMI-AQ participants would see videos of movements performed by either the male or female model, based on whichever model they self-selected as being the closest match to their own gender identity, and performing with the same hand dominance as they reported for the arm movement items. Definitions of imagery generation and imagery maintenance were provided to participants within the AOMI-AQ. Generation and maintenance were described as the following; "how easily can you create an imagined feeling of movement execution whilst watching the video", and "how easily can you hold the imagined feeling of movement execution throughout the duration of the video", respectively. Completion of the demographic questions and three questionnaires took approximately 30 minutes.

Data analysis

Analysis for Study 1 consisted of two phases: 1) parallel analysis and exploratory factor analysis on the 16-item AOMI-AQ data, and 2) validation of the AOMI-AQ against the KI sub-scales of MIQ-3 and VMIQ-2 questionnaires. For Phase 1, AOMI-AQ assessment was comprised of Horn's parallel analysis (PA) and exploratory factor analysis (EFA). Use of PA, alongside scrutiny of the scree plot, facilitated initial assumptions regarding underlying factor structure. In addition, PA is an empirically supported technique for determining the quantity of factors (Pallant, 2007). For PA, random data sampling was employed (O'Connor, 2000). EFA, using Principal Axis Factoring, verified the number of factors and provided information on the adequacy of underlying data and correlation matrix (Kaiser-Meyer-Olkin, KMO, and Bartlett's Sphericity). For Phase 2, validation of the AOMI-AQ was achieved through Pearson's correlations between Generate scores for the AOMI-AQ and overall scores of the VMIQ-2 and MIQ-3. All tests were conducted with an *a priori* alpha of $p < 0.05$ as a threshold of significance.

Results

PA (using 1000 resamples) indicated that a single factor obtained an eigenvalue greater than random data (i.e., 9.48 vs. 0.72). Scrutiny of the scree plot supported the conclusion that one factor existed. Moreover, EFA revealed the existence of one factor (eigenvalue of 9.41), which explained 61.32% of the variance within the data set. A satisfactory correlation matrix and suitable sampling adequacy emerged, Bartlett's Sphericity, $p < 0.001$ and KMO = 0.95. All items loaded above the strict threshold of 0.6 by Hair et al., (2006). Loadings ranged between 0.70 and 0.87, with an average of 0.77.

It was concluded that participants were not distinguishing between the generate and maintain items, given the support for a single factor and the existence of similar factor loadings for these items (i.e., average loading for generate items = 0.77, average loading for maintain items = 0.76). Reanalysis, using EFA, assessed the legitimacy of an 8-item measure comprising the Generate items only, the subscale which related most strongly to the MIQ-3 questionnaire. PA revealed that a single factor comprised an eigenvalue greater than random data (4.74 vs. 0.45). EFA supported this finding, as one factor emerged with an eigenvalue of 5.14, explaining 64.2% of the variance with the data set. A suitable correlation matrix and sampling adequacy existed (Bartlett's Sphericity, $p < 0.001$, and KMO = 0.89). Factor loadings (Table 1) were greater than 0.6, ranging from 0.71 to 0.83 (average loading was 0.77).

Table 1. Factor loadings for the 8-item AOMI-AQ consisting of only the Generate subscale.

Item	Loading
Knee (first person)	0.83
Jump (first person)	0.80
Knee (third person)	0.80
Toe touch (third person)	0.79
Toe touch (first person)	0.76
Jump (third person)	0.75
Arm (first person)	0.73
Arm (third person)	0.71

Note. Items are presented in descending order, from higher to lower loading values. Factor loadings from EFA derived using Principal Axis Factoring. Loadings represent the relationships between latent and observed variables (i.e., items)

Cronbach's alpha was good-to-excellent. Across the 8-item AOMI-AQ there was an excellent internal consistency (α = 0.92), which was similar for the 12 item VMIQ-2 (α = 0.95). Across the 4 items for the MIQ-3 internal consistency was good-to-excellent (α = 0.893). As illustrated in Figure 2, Pearson's correlation between the generate subscale of the AOMI-AQ and kinesthetic components of the MIQ-3 and VMIQ-2 revealed significant positive correlations. As expected, there was a significant positive correlation between the AOMI-AQ and the MIQ-3 ($r = 0.71$, $p < 0.001$) sharing 42% variance with 58% variance unaccounted. Furthermore, significant positive correlations were also found between the AOMI-AQ and VMIQ-2 ($r = 0.43$, $p < 0.001$), and between VMIQ-2 and MIQ-3 ($r = 0.53$, $p < 0.001$).

Figure 2. Correlations and overlap of responses for the AOMI-AQ Generate subscale and kinesthetic imagery subscales of the MIQ-3 and VMIQ-2. Panels A-C show Pearson's correlations between participants' responses for the three questionnaires. Grey bands represent 95% confidence intervals. The density plot (Panel D) illustrates the similarities in responses between the three questionnaires. Scales were standardized using min-max normalization allowing comparison of the questionnaire responses.

Discussion

Study 1 developed a self-report instrument to measure the ability to perform AOMI. To this end, a video-based questionnaire was adapted from the MIQ-3, which is widely used to assess MI ability. The parallel analysis and exploratory factor analysis identified one underlying factor to the AOMI-AQ, specifying that participants did not differ in their responses to the Generate or Maintain subscales of the questionnaire. Although this could be due to no general differences within our sample between these two dimensions of imagery, a more likely explanation is that participants were not able to distinguish between these two different – but closely related – imagery processes. Thus, we restricted the AOMI-AQ to contain only the Generate items, reducing it to an 8-item questionnaire in total. This decision to retain the Generate items rather than the Maintain items was taken to ensure consistency with the MIQ-3, which measures ease of image generation. Reduction of the AOMI-AQ to eight items also had the benefit of halving the questionnaire content, which makes it more feasible for use by researchers and practitioners. The 8-item AOMI-AQ had an excellent internal consistency, and EFA indicated the questionnaire to explain 64.2% of the data. Furthermore, all factor loadings were above 0.6 indicating all items to be closely related to the underlying factor.

The AOMI-AQ, VMIQ-2 and MIQ-3 correlated positively. As expected the strongest relationship was between the AOMI-AQ and MIQ-3 ($r = 0.71$), which could be interpreted as a relatively strong positive correlation (Gignac & Szodorai, 2016). Both these questionnaires required kinesthetic imagery of the same movements, either with or without simultaneous action observation, and so it is intuitive that there would be a relatively strong correlation between these two measures. Despite these resemblances, however, the correlation coefficient of 0.71 between these two questionnaires may provide indirect evidence that they indeed measure slightly different processes and demonstrates the need for a specific measure of AOMI ability. Specifically, the AOMI-AQ and MIQ-3 account for 42% of shared variance, with 58% variance unaccounted. Responses to the AOMI-AQ and the VMIQ-2 were also positively and significantly correlated, although the correlation was weaker than that between the AOMI-AQ and MIQ-3. This weaker correlation was expected, however, given that the two questionnaires require KI of different actions. In the case of the VMIQ-2, some of the imagined actions may have been unfamiliar to participants (e.g., swinging on a rope or jumping off a high wall), potentially causing difficulties in generating KI of the feeling of those movements due to their lack of recent motor experience performing similar actions (see Olsson & Nyberg, 2010; Olsson & Nyberg, 2011), and resulting in lower KI vividness ratings for the VMIQ-2 (see Figure 2 Panel D).

Finally, correlation between the MIQ-3 and VMIQ-2 shows a positive relationship. This is somewhat congruent with previous literature which compared responses to these questionnaires, where a stronger positive correlation between the KI subscales of the MIQ-3 and VMIQ-2 has been reported ($r = 0.706$; Williams et al., 2012), than was found in the current study ($r = 0.53$). One explanation for the slight disparity between these two studies could be due to the participants recruited; for instance, Williams et al. (2012) recruited student athletes whereas our sample comprised a diverse and non-specific population. In response to training, athletes might acquire more sophisticated motor repertoires which may better lend themselves to the tasks imagined in the VMIQ-2 and could have resulted in the higher correlations being reported by Williams et al. (2012). Furthermore, athletes may also engage more frequently in MI as an adjunct to physical training than participants in our sample. Consequently, this imagery practice may contribute to a better imagery ability and ratings across the MIQ-3 and VMIQ-2 in Williams et al.'s sample. It has been suggested that independent MI would be more effective and better leveraged to improve task-specific performance in experts compared to in novices or less experienced individuals, primarily due to experts having more physical experience to inform their imagery (Zhang et al., 2018). AOMI, however, has been proposed to be similarly beneficial for both experts and those with less experience (McNeill et al., 2020), and so perhaps similar AOMI-AQ responses may be expected across these populations. This, however, should be determined through future investigations researching expertise dependent differences in AOMI engagement and performance benefits.

Study 2

As reported in Study 1, the AOMI-AQ was found to be a robust measure of AOMI ability. Since exploratory factor analysis indicated the presence of one underlying factor, an informed decision was taken to remove the maintain items from the AOMI-AQ, leaving an 8-item measure focusing on measurement of AOMI generation ability. Study 2 therefore involved further investigation of the modified 8-item AOMI-AQ measure. A confirmatory factor analysis on the 8-item AOMI-AQ was first conducted, before establishing the test-retest reliability of the AOMI-AQ. Finally, an exploratory analysis comparing scores between researcher supervised and unsupervised completions of the AOMI-AQ was conducted to gain insight into whether researcher supervision was a necessary requirement when administering the AOMI-AQ. In practice the AOMI-AQ will ideally be administered in the presence of a researcher, educator, or coach; however, it may also be convenient to be able to deliver this measure remotely without supervision.

Methods

Participants

Participants who completed Study 1 were invited to take part in Study 2, and 174 volunteered. The sample comprised 106 females and 68 males. The mean age of the sample was 30.1 years (*SD* = 9.12; range = 18-57 years). There were 154 right-handed participants and 20 were left-handed. According to Park et al. (2018), a sample of 174 participants for this study is sufficient for accurate test-retest reliability assessment of the AOMI-AQ. Furthermore, the current sample exceeded more recent reliability assessments of imagery ability questionnaires, such as the MIQ-3 (Suárez Rozo et al., 2022; Yunus et al., 2021) and VMIQ-2 (Plakoutsis et al., 2023; Ziv et al., 2017).

Materials

Participants completed the 8-item AOMI-AQ. This required participants to observe videos of four movements (a toe touch, crouch and jump, arm raise, and knee raise) from two visual perspectives (first- and third-person), whilst simultaneously imagining the feeling of executing the observed movements. Participants then rated the difficulty/ease of their KI generation on a 7-point Likert type (1 – "*Very hard to feel*" to 7 – "*Very easy to feel*"). Full details of the AOMI-AQ are reported in Study 1.

Procedures

Study 2 involved participants completing the AOMI-AQ for a second time, within 7-10 days of the first testing session (*M* duration = 8.02 days, *SD* = 2.3). This 7-10 day period was similar to that used in recent test-retest reliability protocols for other imagery ability measures (Plakoutsis et al., 2023; Suárez Rozo et al., 2022). As shown in Figure 3, for this second session participants were randomly assigned to complete the AOMI-AQ with supervision from a researcher (*n*=82) or without supervision (*n*=92). Supervised testing sessions followed a similar format to Study 1, in which a researcher led the testing session and ensured that participants physically performed the movements instructed in the AOMI-AQ when required. For unsupervised testing sessions, participants were sent a link to the AOMI-AQ by email and instructed to complete the questionnaire independently. This aspect allowed confirmation of whether the self-administration of the questionnaire would be reliable and appropriate for future use. To encourage adherence to the instructions, the self-administered version of the AOMI-AQ included a fixed response delay of 10 seconds during physical performance and AOMI components of the questionnaire to ensure that participants could not continue until sufficient time to perform the actions when required had elapsed. Response times when providing AOMI ratings were also recorded to ensure the video was observed in full. Based on reaction times we checked to ensure adequate time was taken to have performed or imagined the task.

Figure 3. Study 1 and Study 2 timelines and the involvement of participants at each stage of the AOMI-AQ development. All participants completed the initial AOMI-AQ assessment and completed the MIQ-3 and VMIQ-2 for validation of the AOMI-AQ. Of these participants, 174 completed the retest AOMI-AQ which was either completed with the supervision of a researcher or independently.

Data analysis

Confirmatory factor analysis was used to verify the factor solution from Study 1. The weighted least square mean and variance adjusted estimation method calculated model fit and parameter estimates. This is suitable for data that contains ordinal characteristics (Li, 2016). Fit indices of chi-square, Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Standardized Root-Mean-Square Residual (SRMR) were included when evaluating data-model fit. Good fit thresholds were CFI ≥ 0.95, TLI ≥ 0.95, and SRMR ≤ 0.08 (Hu & Bentler, 1995). Alongside chi-square, the normed chi-square was presented given the sensitivity of chi-square to sample size, with values <1 suggestive of overfitting (Obst & White, 2005).

To establish the test-retest reliability of the AOMI-AQ we conducted a Pearson's correlation between participants' initial completion of the AOMI-AQ and the retest dataset. In addition, an independent-samples t-test was conducted comparing the scores of participants who completed the retest with the supervision of a researcher and without supervision. We also analyzed potential differences between participants' initial supervised AOMI-AQ responses with their supervised or unsupervised retests using paired-samples t-tests.

Results

Mean values for the completions of the supervised and unsupervised AOMI-AQ are presented in Table 2. Confirmatory factor analysis revealed good fit for the hypothesised unidimensional model, χ^2 (20) = 73.21, $p < 0.001$, normed $\chi^2 = 3.66$, CFI = 0.99, TLI = 0.98, SRMR = 0.02. Scrutiny of factor loadings indicated that all items loaded above 0.6, thus meeting the strict requirements of Hair et al. (1998). Item loadings ranged from 0.75 to 0.93, with an average loading of 0.85. This supported the existence of a single dimension underpinning the measure, and findings aligned with the emergence of a single dimension within Study 1.

Table 2. Means (M) and standard deviation (SD) for responses for the test-retest of the AOMI-AQ when supervised and unsupervised.

Test	Supervision	*N*	*M*	*SD*
AOMI-AQ	Supervised	174	45.02	8.35
AOMI-AQ retest	Supervised	82	45.12	8.94
	Unsupervised	92	44.79	8.49
	Overall	174	44.95	8.68

Cronbach's Alpha for the retest dataset indicated an excellent internal reliability for the AOMI-AQ ($\alpha = 0.941$), which was similar to the Study 1 dataset ($\alpha = 0.92$). Correlational analysis between these two datasets showed there to be a significant positive correlation ($r = 0.74$, $p < 0.001$). Individuals who rated their AOMI ability highly on the initial test also reported high scores at retest, and the reverse was true for those who rated lower (see Figure 4).

Pearson's correlations between the initial supervised AOMI-AQ and retest AOMI-AQ indicated significant positive correlations for the supervised ($r = 0.83$, $p < 0.001$) and unsupervised participants (r

= 0.65, $p < 0.001$). An independent samples t-test comparing retest responses for individuals who were supervised ($M = 44.94$, $SD = 8.27$) to those who were unsupervised ($M = 44.95$, $SD = 8.68$) indicated no differences in responses, $t(167.35) = 0.248$, $p = 0.805$. Furthermore, paired-samples t-tests revealed no statistical differences between participants' initial AOMI-AQ responses and their retest response depending on whether they were supervised, $t(81) = 0.131$, $p = 0.896$, or unsupervised, $t(169.78) = 0.312$, $p = 0.756$.

Figure 4. Panel A illustrates the similarities between participants' initial completion of the AOMI-AQ and their retest which was completed 7-10 days later. The correlation between these two datasets is represented in Panel B with the grey band representing 95% confidence intervals.

Discussion

The aim of Study 2 was twofold: (i) to establish the test-retest reliability of the AOMI-AQ, and (ii) to establish whether differences in AOMI-AQ scores exist when completed under supervision from a researcher or independently by participants. Confirmatory factor analysis confirmed good factor loadings for each item, and in agreement with Study 1, confirmed only one underlying factor. The test-retest reliability findings demonstrate that the AOMI-AQ, when combining supervised and unsupervised retests, has acceptable-to-good test-retest reliability (Cicchetti, 1994; Nunnally, 1978), indicated by a significant positive correlation between test-retest datasets ($r = 0.74$). Isolating these responses to only individuals who were supervised for both testing sessions showed excellent test-retest reliability ($r = 0.83$; Cicchetti, 1994; Nunnally, 1978). Evidence that the AOMI-AQ has good-to-excellent test-retest reliability, in which participants' scores on the questionnaire remain stable across testing sessions, enhances the possible applications of the AOMI-AQ in research and applied settings. Specifically, this finding opens up the possibility to use the AOMI-AQ as an outcome measure for researchers or applied practitioners seeking to establish techniques to improve AOMI abilities of participants or clients.

Furthermore, there were no statistical differences between participants who completed the retest under the supervision of a researcher or independently without supervision. Both supervised and unsupervised datasets were positively correlated to participants' initial AOMI-AQ datasets. However, the correlation for unsupervised participants was weaker ($r = 0.65$), which could be interpreted as a fair-to-good retest reliability (Cicchetti, 1994; Nunnally, 1978). Given that time constraints are often a factor when conducting both laboratory-based and applied AOMI research, this finding provides tentative reassurance that the AOMI-AQ can be administered to individuals remotely and without supervision, although supervised completion is recommended where possible.

General discussion

No measure to quantify AOMI ability existed prior to the completion of these two studies. To address this gap in the literature, we developed and validated a tool to accurately assess an individual's AOMI ability. Study 1 established a measure of AOMI ability (i.e., the AOMI-AQ) and validated this measure against previously established MI questionnaires. Study 2 sought to determine the test-retest reliability of the AOMI-AQ and establish whether this measure could be completed without researcher supervision.

Collectively, the results indicate that the AOMI-AQ is a valid and reliable measure of participants' ability to generate KI during congruent action observation. Study 1 showed the original 16-item AOMI-AQ to have one underlying factor, indicating that participants did not distinguish between Generation and Maintenance subscales. Consequently, the maintenance subscale was removed to create an 8-item version of the AOMI-AQ focused on measuring the ability of participants to generate KI during concurrent action observation. The 8-item AOMI-AQ correlated positively with the kinesthetic imagery subscales of both the MIQ-3 and VMIQ-2, both of which are valid measures of imagery ability (Williams et al., 2012; Roberts et al., 2008). Study 2 then established that the AOMI-AQ has good-to-excellent test-retest reliability and demonstrated that moderate/fair AOMI-AQ test-retest reliability may be obtained when completed independently.

Prior to the development of the AOMI-AQ, in the absence of an appropriate measure, researchers investigating AOMI have previously attempted to quantify AOMI ability using pre-existing MI ability questionnaires. Whilst these undoubtedly measure processes related to AOMI, it is recognised the ability to generate MI involves different processes to those required for AOMI (Scott et al., 2022). A major distinction between conventional uses of MI and AOMI is the presence of a visual display during AOMI. This display provides visual content which the individual then uses as a scaffold and stimulus to generate their kinesthetic representation of the observed action (Eaves et al., 2022), which could also be considered an explicit form of what has been referred to as 'cross-modal imagery' (Nanay, 2018; Spence & Deroy, 2012). While the validation of the AOMI-AQ supports its likeness to the MIQ-3 and VMIQ-2, the discovery of only moderate-to-large positive correlations between the AOMI-AQ and MIQ-3 (r = 0.71) and VMIQ-2 (r = 0.43) may provide indirect support that different processes were measured, providing tangential justification for a specific measure of AOMI ability.

Excellent test-retest reliability for the supervised AOMI-AQ provides reassurance that the AOMI-AQ can be administered over time and interpreted with confidence. The unsupervised retest responses, however, should be interpreted with caution depending on the mode of delivery. Pearson's correlation for the unsupervised responses, while positively correlated to initial supervised response, showed a weaker correlation than the supervised datasets (rs = 0.65 and 0.83, respectively). Although the unsupervised retest was interpreted as having fair-to-good retest reliability (Cicchetti, 1994; Nunnally, 1978), one consideration regarding these results is that administering the AOMI-AQ unsupervised after a supervised assessment may result in slight variations in responses. An alternative interpretation of this finding could be that the mode of delivery of the AOMI-AQ should be kept consistent across measurements (i.e., always supervised or independently), to ensure responses are most comparable across measurements. Nevertheless, consistent supervised delivery of this tool should produce highly reliable responses when monitoring AOMI ability over time.

These findings may inform previous assumptions regarding the multidimensional nature of AOMI, that is, the requirements to generate, maintain and transform a kinesthetic representation (Scott et al., 2022). Reference to imagery dimensions such as the generation, maintenance, inspection and transformation of content in MI and AOMI was adopted from Kosslynian frameworks of VI (Cumming &

Eaves, 2018; Dror & Kosslyn, 1994). In Study 1, participants were seemingly unable to distinguish between the concepts of image generation and image maintenance. As this questionnaire was developed based on the MIQ-3, which measures imagery generation abilities, we propose that the AOMI-AQ was accurate in the measurement of this aspect of imagery. While it could be the case that AOMI is not as multidimensional as VI or MI (e.g., Cumming & Eaves, 2018; Dror & Kosslyn, 1994; Kraeutner et al., 2020), and requirements for VI are indeed limited during AOMI (Wright et al., 2022), it may also be that subjective based measures such as self-report questionnaires are not sensitive to the maintenance property of AOMI. Therefore, this aspect of AOMI may best be captured and quantified through other methods such as neurophysiological or chronometry-based measures.

The AOMI-AQ provides the first valid and reliable tool by which AOMI ability can be quantified and this has multiple benefits for both research and applied practice. For example, in research contexts, the ability to accurately measure AOMI ability now provides researchers with a tool to (i) introduce participants to the concept of AOMI, (ii) screen for AOMI ability as part of study inclusion/exclusion criteria, (iii) control for AOMI ability or split participants based on AOMI ability when allocating experimental groups, (iv) monitor changes in AOMI ability as an outcome measure in research, and (v) report participant AOMI abilities (Moreno-Verdú et al., 2024), without having to rely on MI ability measures as a proxy measure for AOMI. Similarly, in applied contexts, practitioners can now assess and monitor AOMI ability prior to and throughout AOMI training programmes and interventions administered to their clients. Moreover, in applied settings, the AOMI-AQ could be used alongside previously established MI and AO ability measures to help determine which simulation approach may be easier, more engaging, or better suited to their client, allowing them to tailor simulation interventions based on their clients needs and preferences. In all these cases researchers and practitioners may find it helpful to consider cut-off values for distinguishing ‘good’ and ‘poor’ AOMI ability. While future research may help establish such values, appropriate initial criteria based on the recommendations of Robin and Coudevylle (2018) and Robin and Blandin (2021) could be to consider mean AOMI-AQ scores > 5 and < 2, respectively, as indicative of good and poor AOMI ability.

A potential limitation to this research is the limited application which the AOMI-AQ may have for specific populations who may have physical impairments or differences, which make our models less relatable. Although the requirement to physically execute the four movements before engaging in AOMI served to provide participants with recent motor experience on which to base their KI generation, the measure may not be suitable for use with certain populations whose movement abilities may be impaired. For example, in sport and clinical contexts, athletes with certain physical disabilities or injuries, or individuals who have experienced stroke or other clinical motor impairment, may be unable to execute the actions required to complete the AOMI-AQ. Future iterations of the AOMI-AQ should, therefore, consider the use of models with of varied movement capabilities such as clinical populations – for whom AOMI has shown to be an effective intervention (Bek et al., 2021; Binks et al., 2023; Scott et al., 2023; Sun et al., 2016) – and also the use of different tasks, similar to the Movement Imagery Questionnaire – Revised second edition (Greg et al., 2010). Similarly, we did not isolate and assess AOMI-AQ responses in athletes, as previous imagery questionnaires have done (Williams et al., 2012). It would be beneficial to determine whether individuals with well developed motor repertoires and those with less experience would respond differently to the AOMI-AQ, as it has been proposed that AOMI could be beneficial for both experts and novices (McNeill et al., 2020).

While comparable to the VMIQ-2 and MIQ-3, the AOMI-AQ had a high Cronbach alpha score across the two studies (α = 0.92-0.941). Although these scores are indicative of a high internal consistency, this may also suggest similarities between the items. Therefore, future research should focus on refinement of the AOMI-AQ to ensure efficiency in its delivery. In addition, it is important to note that the AOMI-AQ was established to measure the ability to observe and imagine the same action simultaneously; often referred to as congruent AOMI (Eaves et al., 2022; Vogt et al., 2013). The current tool has not been validated for less commonly used alternative forms of AOMI, such as coordinative or conflicting AOMI, where the simultaneous imagery and observation content differ from each other to varying extents (Vogt et al., 2013).

To conclude, the present studies demonstrate the newly developed AOMI-AQ to be a valid and reliable measure of AOMI ability. This new tool should advance future AOMI research and applied practice by providing a direct measure of AOMI ability, negating the current reliance of AOMI researchers and practitioners on less appropriate MI-based measures as a proxy measure of AOMI ability. It has been proposed that AOMI is more beneficial than independent AO and may have theoretical and practical advantages over traditional uses of MI (Chye et al., 2022; Scott et al., 2022). Accordingly, there has been increased interest in the application of AOMI across sport and rehabilitation settings to improve behavioural outcomes. The AOMI-AQ provides researchers and coaches who choose to implement AOMI interventions with a reliable tool to assess an individual's ability to use AOMI and monitor changes in AOMI ability over time and across interventions or training periods.

References

Binks, J. A., Emerson, J. R., Scott, M. W., Wilson, C., Van Schaik, P., & Eaves, D. L. (2023). Enhancing upper-limb neurorehabilitation in chronic stroke survivors using combined action observation and motor imagery therapy. *Frontiers in Neurology*, *14*, 1097422. https://doi.org/10.3389/fneur.2023.1097422

Chye, S., Valappil, A. C., Wright, D. J., Frank, C., Shearer, D. A., Tyler, C. J., ... & Bruton, A. M. (2022). The effects of combined action observation and motor imagery on corticospinal excitability and movement outcomes: Two meta-analyses. *Neuroscience & Biobehavioral Reviews*, *143*, 104911. https://doi.org/10.1016/j.neubiorev.2022.104911

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

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/10.1177/0276236617752439

Dance, C. J., Ward, J., & Simner, J. (2021). What is the link between mental imagery and sensory sensitivity? Insights from aphantasia. *Perception*, *50*(9), 757–782. https://doi.org/10.1177/03010066211042186

De Vries, S., Tepper, M., Feenstra, W., Oosterveld, H., Boonstra, A. M., & Otten, B. (2013). Motor imagery ability in stroke patients: the relationship between implicit and explicit motor imagery measures. *Frontiers in human neuroscience, 7*, 790. https://doi.org/10.3389/fnhum.2013.00790

Dilek, B., Ayhan, C., & Yakut, Y. (2020). Reliability and validity of the Turkish version of the movement imagery questionnaire-3: Its cultural adaptation and psychometric properties. Neurological *Sciences and Neurophysiology, 37*(4), 221. https://doi.org/10.4103/NSN.NSN_30_20

Dror, I. E., & Kosslyn, S. M. (1994). Mental imagery and aging. *Psychology and aging*, 9(1), 90. https://doi.org/10.1037/0882-7974.9.1.90

Eaves, D. L., Riach, M., Holmes, P. S., & Wright, D. J. (2016a). 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., Behmer Jr, L. P., & Vogt, S. (2016b). 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., Hodges, N. J., Buckingham, G., Buccino, G., & Vogt, S. (2022). Enhancing motor imagery practice using synchronous action observation. Psychological Research, 1-17. https://doi.org/10.1007/s00426-022-01768-7

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

Emerson, J. R., Scott, M. W., Van Schaik, P., Butcher, N., Kenny, R. P., & Eaves, D. L. (2022). A neural signature for combined action observation and motor imagery? An fNIRS study into prefrontal activation, automatic imitation, and self–other perceptions. *Brain and Behavior*, *12*(2), e2407. https://doi.org/10.1002/brb3.2407

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. Personality and individual differences, 102, 74-78. https://doi.org/10.1016/j.paid.2016.06.069

Gregg, M., Hall, C., & Butler, A. (2010). The MIQ-RS: a suitable option for examining movement imagery ability. *Evidence-Based Complementary and Alternative Medicine, 7*, 249-257. https://doi.org/10.1093/ecam/nem170

Gulyás, E., Gombos, F., Sütöri, S., Lovas, A., Ziman, G., & Kovács, I. (2022). Visual imagery vividness declines across the lifespan. *Cortex*, 154, 365–374. https://doi.org/10.1016/j.cortex.2022.06.011

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis* (5th ed.). Prentice Hall.

Hardy, L., & Callow, N. (1999). Efficacy of external and internal visual imagery perspectives for the enhancement of performance on tasks in which form is important. *Journal of Sport and Exercise Psychology*, *21*(2), 95–112. https://doi.org/10.1123/jsep.21.2.95

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

Hu, L., & Bentler, P. (1995). *Measuring model fit. Structural equation modeling: Concepts, issues and applications.* Sage.

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

Kraeutner, S. N., Eppler, S. N., Stratas, A., & Boe, S. G. (2020). Generate, maintain, manipulate? Exploring the multidimensional nature of motor imagery. *Psychology of Sport and Exercise*, *48*, 101673. https://doi.org/10.1016/j.psychsport.2020.101673

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2010). Multimodal images in the brain. In A. Guillot & C. Collet (Eds.), The Neurophysiological Foundations of Mental and Motor Imagery (pp. 3–16). Oxford University Press.

La Touche, R., Grande-Alonso, M., Cuenca-Martínez, F., Gónzález-Ferrero, L., Suso-Martí, L., & Paris-Alemany, A. (2019). Diminished kinesthetic and visual motor imagery ability in adults with chronic low back pain. *PM & R: The Journal of Injury, Function, and Rehabilitation*, *11*(3), 227–235. https://doi.org/10.1016/j.pmrj.2018.05.025

Malouin, F., Richards, C. L., Durand, A., Descent, M., Poiré, D., Frémont, P., ... & Doyon, J. (2009). Effects of practice, visual loss, limb amputation, and disuse on motor imagery vividness.

Neurorehabilitation and Neural Repair, *23*(5), 449–463. https://doi.org/10.1177/1545968308328733

Marshall, B., Wright, D. J., Holmes, P. S., 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

Martini, R., Carter, M. J., Yoxon, E., Cumming, J., & Ste-Marie, D. M. (2016). Development and validation of the Movement Imagery Questionnaire for Children (MIQ-C). *Psychology of Sport and Exercise, 22*, 190-201. https://doi.org/10.1016/j.psychsport.2015.08.008

McNeill, E., Ramsbottom, N., Toth, A., & Campbell, M. (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

McNeill, E., Toth, A. J., Harrison, A. J., & Campbell, M. J. (2020). Cognitive to physical performance: a conceptual model for the role of motor simulation in performance. *International Review of Sport and Exercise Psychology, 13*(1), 205-230. https://doi.org/10.1080/1750984X.2019.1689573

McNeill, E., Toth, A. J., Ramsbottom, N., & Campbell, M. J. (2021). Self-modelled versus skilled-peer modelled AO+ MI effects on skilled sensorimotor performance: A stage 2 registered report. *Psychology of Sport and Exercise*, *54*, 101910. https://doi.org/10.1016/j.psychsport.2021.101910

Moreno-Verdú, M., Hamoline, G., Van Caenegem, E. E., Waltzing, B. M., Forest, S., Valappil, A. C., ... & Hardwick, R. M. (2023). Guidelines for Reporting Action Simulation Studies (GRASS): proposals to improve reporting of research in Motor Imagery and Action Observation. Neuropsychologia, 108733. https://doi.org/10.1016/j.neuropsychologia.2023.108733

Morris, T., Spittle, M., & Watt, A. P. (2005). *Imagery in sport*. Human Kinetics.

Nanay, B. (2018). Multimodal mental imagery. *Cortex*, *105*, 125–134. https://doi.org/10.1016/j.cortex.2017.07.006

Nunnally, J. C. (1978). *Psychometric theory. 2nd edition*. McGraw-Hill; New York: 1978.

Obst, P., & White, K. (2005). Three-dimensional strength of identification across group memberships: A confirmatory factor analysis. *Self and Identity*, *4*(1), 69-80. https://doi.org/10.1080/13576500444000182

Olsson, C. J., & Nyberg, L. (2010). Motor imagery: if you can't do it, you won't think it. *Scandinavian journal of medicine & science in sports, 20*(5), 711-715. https://doi.org/10.1111/j.1600-0838.2010.01101.x

Olsson, C. J., & Nyberg, L. (2011). Brain simulation of action may be grounded in physical experience. *Neurocase, 17*(6), 501-505. https://doi.org/10.1080/13554794.2010.547504

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, and Computers, 32,* 396–402. https://doi.org/10.3758/BF03200807

O'Shea, H. (2022). Mapping relational links between motor imagery, action observation, action-related language, and action execution. *Frontiers in Human Neuroscience*, *16*. https://doi.org/10.3389/fnhum.2022.984053

Park, M. S., Kang, K. J., Jang, S. J., Lee, J. Y., & Chang, S. J. (2018). Evaluating test-retest reliability in patient-reported outcome measures for older people: A systematic review. *International Journal of Nursing Studies*, *79*, 58–69. https://doi.org/10.1016/j.ijnurstu.2017.11.003

Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis for Windows* (3rd ed.). Open University Press.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-01193-y

Plakoutsis, G., Fousekis, K., Tsepis, E., & Papandreou, M. (2023). Cross cultural adaptation, validity and reliability of the Greek version of the Vividness of Movement Imagery Questionnaire-2 (VMIQ-2). *Discover Psychology, 3*(1), 30. https://doi.org/10.1007/s44202-023-00091-5

Roberts, R., Callow, N., Hardy, L., Markland, D., & Bringer, J. (2008). Movement imagery ability: development and assessment of a revised version of the vividness of movement imagery questionnaire. *Journal of Sport and Exercise Psychology*, *30*(2), 200–221. https://doi.org/10.1123/jsep.30.2.200

Robin, N., & Blandin, Y. (2021). Imagery ability classification: Commentary on «kinaesthetic imagery ability moderates the effect of an AO+MI intervention on golf putt performance: A pilot study» by McNeill, Ramsbottom, Toth, and Campbell (2020). *Psychology of Sport and Exercise, 57*, 102030. https://doi.org/10.1016/j.psychsport.2021.102030

Robin, N., & Coudevylle, G. R. (2018). The influences of tropical climate, imagery ability, distance and load on walking time. *International Journal of Sport Psychology, 3*, 66-87. https://doi.org/10.1080/20445911.2017.1384382

Romano-Smith, S., Roberts, J. W., Wood, G., Coyles, G., & Wakefield, C. J. (2022). Simultaneous and alternate combinations of action-observation and motor imagery involve a common lower-level sensorimotor process. *Psychology of Sport and Exercise*, *63*, 102275. https://doi.org/10.1016/j.psychsport.2022.102275

Rungsirisilp, N., & Wongsawat, Y. (2022). Applying Combined Action Observation and Motor Imagery to Enhance Classification Performance in a Brain–Computer Interface System for Stroke Patients. *IEEE Access*, *10*, 73145-73155. https://doi.org/10.1109/ACCESS.2022.3190798

Scott, M., Taylor, S., Chesterton, P., Vogt, S., & Eaves, D. L. (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/10.1016/j.ajsep.2022.07.002

Scott, M. W., Wright, D. J., Smith, D., & Holmes, P. S. (2022). Twenty years of PETTLEP imagery: An update and new direction for simulation-based training. *Asian Journal of Sport and Exercise Psychology*, *2*(2), 70–79. https://doi.org/10.1016/j.ajsep.2022.07.002

Scott, M. W., Wood, G., Holmes, P. S., Marshall, B., Williams, J., & Wright, D. J. (2023). Combined action observation and motor imagery improves learning of activities of daily living in children with

Developmental Coordination Disorder. *Plos one*, *18*(5), e0284086. https://doi.org/10.1371/journal.pone.0284086

Simonsmeier, B. A., Andronie, M., Buecker, S., & Frank, C. (2021). The effects of imagery interventions in sports: A meta-analysis. *International Review of Sport and Exercise Psychology*, *14*(1), 186–207. https://doi.org/10.1080/1750984X.2020.1780627

Spence, C., & Deroy, O. (2013). Crossmodal mental imagery. Multisensory imagery, 157-183. https://doi.org/10.1007/978-1-4614-5879-1_9

Spruijt, S., van der Kamp, J., & Steenbergen, B. (2015). Current insights in the development of children's motor imagery ability. *Frontiers in Psychology*, *6*, 787. https://doi.org/10.3389/fpsyg.2015.00787

Suárez Rozo, M. E., Trapero-Asenjo, S., Pecos-Martín, D., Fernández-Carnero, S., Gallego-Izquierdo, T., Jiménez Rejano, J. J., & Nunez-Nagy, S. (2022). Reliability of the Spanish Version of the Movement Imagery Questionnaire-3 (MIQ-3) and Characteristics of Motor Imagery in Institutionalized Elderly People. *Journal of Clinical Medicine, 11*(20), 6076. https://doi.org/10.3390/jcm11206076

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. M., Annoni, J. M., & 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. J., McNeill, E., Hayes, K., Moran, A. P., & 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

Trapero-Asenjo, S., Gallego-Izquierdo, T., Pecos-Martín, D., & Nunez-Nagy, S. (2021). Translation, cultural adaptation, and validation of the Spanish version of the Movement Imagery Questionnaire-

3 (MIQ-3). *Musculoskeletal Science and Practice, 51*, 102313. https://doi.org/10.1016/j.msksp.2020.102313

Williams, S. E., Cumming, J., Ntoumanis, N., Nordin-Bates, S. M., Ramsey, R., & Hall, C. (2012). Further validation and development of the movement imagery questionnaire. *Journal of Sport and Exercise Psychology*, *34*(5), 621–646. https://doi.org/10.1123/jsep.34.5.621

Wright, D. J., Frank, C., & Bruton, A. M. (2022). Recommendations for combining action observation and motor imagery interventions in sport. *Journal of Sport Psychology in Action*, *13*(3), 155–167. https://doi.org/10.1080/21520704.2021.1971810

Wright, D. J., McCormick, S. A., Birks, S., Loporto, M., & Holmes, P. S. (2015). Action observation and imagery training improve the ease with which athletes can generate imagery. *Journal of Applied Sport Psychology*, *27*(2), 156–170. https://doi.org/10.1080/10413200.2014.968294

Wright, D. J., Wood, G., Eaves, D. L., Bruton, A. M., Frank, C., & Franklin, Z. C. (2018). 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

Yunus, U. Ğ. U. R., COŞKUN, H., & ŞENYURT, A. Y. (2021). THE MOVEMENT IMAGERY QUESTIONNAIRE-3: RELIABILITY AND VALIDITY STUDY ON TURKISH SAMPLE. SPORMETRE *Beden Eğitimi ve Spor Bilimleri Dergisi, 19*(4), 145-156. https://doi.org/10.33689/spormetre.938945

Zhang, L. L., Pi, Y. L., Shen, C., Zhu, H., Li, X. P., Ni, Z., ... & Wu, Y. (2018). Expertise-level-dependent functionally plastic changes during motor imagery in basketball players. *Neuroscience, 380*, 78-89. https://doi.org/10.1016/j.neuroscience.2018.03.050

Ziv, G., Lidor, R., Arnon, M., & Zeev, A. (2017). The Vividness of Movement Imagery Questionnaire (VMIQ-2)–translation and reliability of a Hebrew version. *Israeli Journal of Psychiatry, 54*, 48-53.