



LJMU Research Online

Scott, MW, Esselaar, M, Dagnall, N, Denovan, A, Marshall, B, Deacon, AS, Holmes, PS and Wright, DJ

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

<http://researchonline.ljmu.ac.uk/id/eprint/23500/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Scott, MW, Esselaar, M, Dagnall, N, Denovan, A, Marshall, B, Deacon, AS, Holmes, PS and Wright, DJ (2024) Development and Validation of the Combined Action Observation and Motor Imagery Ability Questionnaire. Journal of Sport & Exercise Psychology. pp. 1-14. ISSN 0895-2779

LJMU has developed [LJMU Research Online](http://researchonline.ljmu.ac.uk/) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

1 **Development and Validation of the Combined Action Observation and Motor Imagery Ability**

2 **Questionnaire**

3

4 Matthew W. Scott^{1,2}, Maaïke Esselaar³ Neil Dagnall⁴, Andrew Denovan⁵, Ben Marshall³,

5 Aimee S. Deacon⁴, Paul S. Holmes³ & David J. Wright^{4*}

6

7

8

9

10 **Affiliations**

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

18

19

20

21

22 *** Corresponding author**

23 Dr. David J. Wright

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

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

26

27 **Abstract**

28 Combined use of action observation and motor imagery (AOMI) is an increasingly popular motor
29 simulation intervention, which involves observing movements on video whilst simultaneously imagining
30 the feeling of movement execution. Measuring and reporting participant imagery ability characteristics is
31 essential in motor simulation research, but no measure of AOMI ability currently exists. Accordingly, the
32 AOMI ability questionnaire (AOMI-AQ) was developed to address this gap in the literature. In Study 1,
33 211 participants completed the AOMI-AQ and the kinesthetic imagery sub-scales of the Movement
34 Imagery Questionnaire-3 and Vividness of Motor Imagery Questionnaire-2. Following exploratory factor
35 analysis, an 8-item AOMI-AQ was found to correlate positively with existing motor imagery measures. In
36 Study 2, 174 participants completed the AOMI-AQ a second time 7-10 days later. Results indicate a good
37 test-retest reliability for the AOMI-AQ. The new AOMI-AQ measure provides a valid and reliable tool
38 for researchers and practitioners wishing to assess AOMI ability.

39

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

42

43 **Introduction**

44 Motor imagery (MI) is a perceptual-like experience involving the mental generation, manipulation
45 and maintenance of the visual and kinesthetic properties of a movement (Kosslyn et al., 2010). Motor
46 imagery ability, therefore, refers to how well an individual can create, maintain and control their motor
47 imagery (Morris et al., 2005). MI is used widely in psychological intervention programmes to improve
48 physical performance (e.g., Robin et al., 2023; for a meta-analysis see Toth et al., 2020) and enhance
49 psychological processes related to performance (e.g., motivation; Simonsmeier et al., 2021). However,
50 within the last decade, imagery research has turned to alternative methods of simulation-based training.
51 One method which has gained interest is the combined and simultaneous use of action observation and
52 motor imagery (AOMI). AOMI involves watching a video or live display of an action, while generating,
53 maintaining, and transforming a time-synched kinesthetic representation of the same action (Scott et al.,
54 2022). In practice, AOMI therefore involves watching movements whilst imagining concurrently the
55 kinesthetic sensations of executing the observed action. Although instructions are typically limited to
56 kinesthetic imagery during AOMI, this does not exclude the spontaneous use of, or requirement for, visual
57 imagery (VI; see Mizuguchi et al., 2016; Scott et al., 2022); however, the visual display during AOMI
58 reduces the requirement for VI to some extent (Wright et al., 2022).

59 Though some studies almost 20 years ago used an approach resembling AOMI (e.g., Smith &
60 Holmes, 2004), the technique has gained increased attention within the last decade (Eaves et al., 2016a;
61 Vogt et al., 2013). Research into AOMI has focused on exploring its neurophysiological or behavioural
62 effects in relation to independent MI or AO (Eaves et al., 2016a; 2022; O’Shea, 2022; Scott et al., 2021;
63 2022). Recent meta-analytical findings indicate that AOMI elicits increased activity in areas of the brain
64 related to motor planning and execution and produces superior performance outcomes compared to
65 independent AO, and effects are at least equivalent to MI on both outcome measures (Chye et al., 2022).

66 There are several theoretical and practical constraints associated with MI that can be resolved by
67 AOMI. For example, as many as 4% of individuals report being unable to generate VI or experience
68 difficulties in doing so (i.e., Aphantasia; Dance et al., 2022). This issue could potentially be addressed

69 through the provision of visual stimuli on video during AOMI. In addition, due to the covert nature of
70 imagery, practitioners have limited control over their client’s imagery experience (i.e., visual perspective,
71 viewing angle, movement timing, and image maintenance; Holmes & Calmels, 2008). These problems may
72 also be resolved through AOMI since the practitioner can control the visual perspective, viewing angle, and
73 movement timing information through the video stimuli they present to their client (Wright et al., 2022).
74 Furthermore, content related to the task and environment, which may vary across individuals based on
75 previous motor or visual experience (Malouin et al., 2009; Wright et al., 2015) are controlled in AOMI
76 (Scott et al., 2022). Another advantage of AOMI over MI is the opportunity for controlling the agent of the
77 simulated action (i.e., via self- or other-modelling). This is important because some individuals may have
78 a natural preference for imagining themselves or others (Holmes & Calmels, 2008).

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

87 In the absence of an appropriate measure of AOMI ability, much of the research to date has failed
88 to include any assessment of participants’ imagery abilities (e.g., Kawasaki et al., 2018; Marshall et al.,
89 2020; Rungsirisilp & Wongsawat, 2022; Taube et al., 2014; 2015). This is problematic as recent guidelines
90 recommend strongly that a measure of imagery ability should be included when reporting AOMI studies
91 (Moreno-Verdú et al., 2022). Alternatively, some investigators have used existing MI ability questionnaires
92 as a proxy measure of AOMI ability (e.g., Emerson et al., 2022; McNeill et al., 2021; Romano-Smith et al.,
93 2022; Scott et al., 2017; Wright et al., 2018). Although this approach seems intuitive in the absence of an
94 existing measure of AOMI ability, this may not be entirely appropriate as AOMI and MI may rely on

95 relatively different neural substrates and utilize different cognitive processes (O'Shea, 2022; Scott et al.,
96 2022). For instance, AOMI results in greater activations bilaterally across the primary motor cortex and
97 cerebellum and stronger activity in the precuneus than independent MI (Taube et al., 2015). Furthermore,
98 the greater recruitment of the rostral pre-frontal cortex during AOMI may indicate different cognitive
99 requirements for its use (Eaves et al., 2016b; Emerson et al., 2022). Conceptually, having the ability to
100 generate and manipulate visual and kinesthetic imagery may not necessarily translate to being able to
101 generate and maintain a kinesthetic imagery (KI) representation in synchrony with an external video or live
102 demonstration. Indeed, it has been proposed that in contrast to typical deliveries of MI, using AOMI
103 requires attentional switching between internal (KI) and external (AO) components (Eaves et al., 2016b;
104 Emerson et al., 2022). Considering these factors, together with the current research interest in AOMI
105 processes, there is a clear need to develop an appropriate measure of AOMI ability.

106 Though no measures of AOMI ability presently exist, the creation of a new instrument can be
107 informed by existing indices of MI ability. It has been proposed that imagery ability consists of different
108 processes, such as the ability to generate, manipulate or maintain imagined content (Eaves & Cummings,
109 2018; Kraeutner et al., 2020). Several measures of imagery ability have been developed, including implicit
110 (i.e., hand laterality judgement task) and explicit (i.e., mental chronometry or self-report) measures, which
111 may probe these different imagery processes (Kraeutner et al., 2020). However, AOMI versions of
112 measures like the hand laterality judgement task or mental chronometry would not be appropriate, as video
113 stimuli would inherently depict the mental rotation and movement timing information being measured in
114 these respective tasks. Therefore, a self-report measure for AOMI ability is most appropriate and has the
115 added benefit of being easy to administer.

116 One of the most widely used measures of MI ability is the Vividness of Movement Imagery
117 Questionnaire-2 (VMIQ-2; Roberts et al., 2008). The VMIQ-2 instructs individuals to imagine 12 actions
118 from first- and third-person visual perspectives and a separate kinesthetic modality, all of which show
119 excellent internal reliability ($\alpha > 0.93$; Roberts et al., 2008). For each action, participants are required to
120 rate the vividness of their imagined movement on a 5-point Likert scale. As the questionnaire imposes no

121 requirement to execute the movements physically prior to imagining them, the movements imagined could
122 conceivably be either within or beyond the individual's own motor repertoire. Williams et al. (2012)
123 previously highlighted limitations in the delivery of this measure due to the open interpretation of some of
124 its items. One item, for example, requires the individual to imagine kicking a ball in the air. This could,
125 conceivably, introduce variation in responses based on the type of ball or kick imagined, which is potentially
126 biased by an individual's previous motor or visual experiences. Other items instruct people to imagine
127 actions that they may never have performed, such as swinging on a rope or jumping off a high wall. An
128 AOMI version of this questionnaire, with the requirement for participants to rate how easily they can
129 imagine the feeling of an action they may have never performed, would clearly be problematic.

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

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

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

148 **Study 1**

149 **Methods**

150 *Participants*

151 Two-hundred and eleven participants ($F = 121$, $M = 89$, non-binary = 1) aged between 18 to 57
152 years ($M = 29$, $SD = 9.03$), with normal or corrected-to-normal vision, participated in Study 1. The
153 relatively wide age range in the sample was appropriate given the ongoing use of AOMI across younger
154 and older age groups (e.g., Emerson et al., 2022; Kawasaki et al., 2018; Mouthon et al., 2020; see Chye et
155 al., 2022 for meta-analysis). Of this sample, 188 participants self-reported being right-handed, with the
156 remainder identifying as left-hand dominant. All participants self-reported being able to perform the
157 actions required to complete the MIQ-3 and AOMI-AQ. The current sample size was sufficient for
158 exploratory factor analysis based on previous guidelines by Comrey and Lee (1992). Furthermore, the
159 sample size exceeded that used in other studies that have adapted the MIQ-3 for use with children
160 (Martini et al., 2016) or for use in different languages (e.g., Dilek et al., 2020; Trapero-Asenjo et al.,
161 2021). Ethical approval for both studies was granted through the Manchester Metropolitan University
162 Faculty of Health and Education Research Ethics and Governance Committee (Ethical Approval
163 Identification Number: 35170) and participants provided written informed consent prior to participating.

164 *Materials*

165 **Development of the combined action observation and motor imagery ability questionnaire**
166 **(AOMI-AQ).** A Qualtrics and Microsoft PowerPoint version of the AOMI-AQ can be accessed and
167 downloaded via https://osf.io/vbqjw/?view_only=3382b7e43a794ed78ea0c17a17eebe1f. A PsychoPy
168 version can be accessed by emailing the authors. The AOMI-AQ was developed by adapting the MIQ-3
169 into a video-based questionnaire depicting similar actions, recorded from both first- and third-person
170 visual perspectives. The MIQ-3 comprises four movements, which individuals perform and subsequently
171 imagine: a right knee raise, a crouch to jump, a horizontal arm adduction (non-dominant limb), and a

172 waist bend toe touch movement. The same movements were used for the AOMI-AQ although the arm
173 action was modified to include an initial vertical abduction, followed by horizontal adduction. This
174 decision was taken to provide additional visual movement information with which participants could
175 synchronize their imagery during AOMI. A library of videos was created with both a male (Caucasian, 28
176 years old) and female (Caucasian, 27 years old) model. Different sex models were included in agreement
177 with AOMI guidelines (Wright et al., 2022), to allow participants to perform AOMI with the video model
178 who matched their chosen gender identity most closely. Each movement was timed to a metronome and
179 filmed from both first- and third-person visual perspectives and showing the models performing as if
180 right- or left-hand dominant, as required for the arm raise action.

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

194 The AOMI-AQ required participants to report how well they were able to generate KI whilst
195 observing a video of each action. In contrast to the MIQ-3, however, no VI rating scale was included in
196 the AOMI-AQ as VI is not typically instructed during AOMI (Scott et al., 2022; Wright et al., 2022). An
197 additional subscale was included requiring participants to rate how well they were able to maintain the

198 imagined feeling throughout the video. As the questionnaire movements ranged from slower, smoother
199 movements (e.g., a knee raise) to more explosive movements (e.g., a crouch-to-jump), the generation of
200 KI throughout movements would presumably fluctuate due to different muscle engagement, force
201 requirements, and movement durations during these movements. For instance, in the crouch-to-jump
202 movement KI should theoretically be strongest during crouch and extension components, and on landing
203 when kinesthetic feedback should be most salient during actual performance, but with limited KI during
204 the flight phase. Maintenance of KI for this jump movement could therefore be expected to differ from
205 that of the other three movements. The inclusion of a maintenance subscale allowed these potential
206 differences to be quantified.

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

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

226

227

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

235

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

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

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

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

267 ***Procedures***

268 Prior to undertaking the study participants were informed of the study aims and provided written
269 informed consent. Next, participants completed a demographic questionnaire in which they reported their
270 age, gender identity, ethnicity, handedness, and whether a male or female model most closely matched
271 their gender identity. Following this, participants were educated on the concept of imagery with an
272 emphasis on KI. Once participants understood the content they would be required to imagine, they
273 completed the AOMI-AQ, VMIQ-2 and MIQ-3 in the presence and under the supervision of a member of
274 the research team, who supervised the testing session and ensured that participants adhered to the
275 instructions to execute and imagine the movements when required. All questionnaires were delivered

276 through PsychoPy (V2023.1.1, Peirce et al., 2019), displayed from a laptop, and the delivery of these
277 questionnaires was counterbalanced across participants to control order effects.

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

287 ***Data analysis***

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

300 **Results**

301 PA (using 1000 resamples) indicated that a single factor obtained an eigenvalue greater than
 302 random data (i.e., 9.48 vs. 0.72). Scrutiny of the scree plot supported the conclusion that one factor
 303 existed. Moreover, EFA revealed the existence of one factor (eigenvalue of 9.41), which explained
 304 61.32% of the variance within the data set. A satisfactory correlation matrix and suitable sampling
 305 adequacy emerged, Bartlett's Sphericity, $p < 0.001$ and KMO = 0.95. All items loaded above the strict
 306 threshold of 0.6 by Hair et al., (2006). Loadings ranged between 0.70 and 0.87, with an average of 0.77.

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

316

317 **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 |

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

321

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

330

331

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

337

338 **Discussion**

339 Study 1 developed a self-report instrument to measure the ability to perform AOMI. To this end, a video-
340 based questionnaire was adapted from the MIQ-3, which is widely used to assess MI ability. The parallel
341 analysis and exploratory factor analysis identified one underlying factor to the AOMI-AQ, specifying that
342 participants did not differ in their responses to the Generate or Maintain subscales of the questionnaire.

343 Although this could be due to no general differences within our sample between these two dimensions of
344 imagery, a more likely explanation is that participants were not able to distinguish between these two
345 different – but closely related – imagery processes. Thus, we restricted the AOMI-AQ to contain only the
346 Generate items, reducing it to an 8-item questionnaire in total. This decision to retain the Generate items
347 rather than the Maintain items was taken to ensure consistency with the MIQ-3, which measures ease of

348 image generation. Reduction of the AOMI-AQ to eight items also had the benefit of halving the
349 questionnaire content, which makes it more feasible for use by researchers and practitioners. The 8-item
350 AOMI-AQ had an excellent internal consistency, and EFA indicated the questionnaire to explain 64.2%
351 of the data. Furthermore, all factor loadings were above 0.6 indicating all items to be closely related to the
352 underlying factor.

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

369 Finally, correlation between the MIQ-3 and VMIQ-2 shows a positive relationship. This is
370 somewhat congruent with previous literature which compared responses to these questionnaires, where a
371 stronger positive correlation between the KI subscales of the MIQ-3 and VMIQ-2 has been reported ($r =$
372 0.706 ; Williams et al., 2012), than was found in the current study ($r = 0.53$). One explanation for the
373 slight disparity between these two studies could be due to the participants recruited; for instance,

374 Williams et al. (2012) recruited student athletes whereas our sample comprised a diverse and non-specific
375 population. In response to training, athletes might acquire more sophisticated motor repertoires which
376 may better lend themselves to the tasks imagined in the VMIQ-2 and could have resulted in the higher
377 correlations being reported by Williams et al. (2012). Furthermore, athletes may also engage more
378 frequently in MI as an adjunct to physical training than participants in our sample. Consequently, this
379 imagery practice may contribute to a better imagery ability and ratings across the MIQ-3 and VMIQ-2 in
380 Williams et al.'s sample. It has been suggested that independent MI would be more effective and better
381 leveraged to improve task-specific performance in experts compared to in novices or less experienced
382 individuals, primarily due to experts having more physical experience to inform their imagery (Zhang et
383 al., 2018). AOMI, however, has been proposed to be similarly beneficial for both experts and those with
384 less experience (McNeill et al., 2020), and so perhaps similar AOMI-AQ responses may be expected
385 across these populations. This, however, should be determined through future investigations researching
386 expertise dependent differences in AOMI engagement and performance benefits.

387 **Study 2**

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

399 **Methods**

400 ***Participants***

401 Participants who completed Study 1 were invited to take part in Study 2, and 174 volunteered.
402 The sample comprised 106 females and 68 males. The mean age of the sample was 30.1 years ($SD = 9.12$;
403 range = 18-57 years). There were 154 right-handed participants and 20 were left-handed. According to
404 Park et al. (2018), a sample of 174 participants for this study is sufficient for accurate test-retest reliability
405 assessment of the AOMI-AQ. Furthermore, the current sample exceeded more recent reliability
406 assessments of imagery ability questionnaires, such as the MIQ-3 (Suárez Rozo et al., 2022; Yunus et al.,
407 2021) and VMIQ-2 (Plakoutsis et al., 2023; Ziv et al., 2017).

408 ***Materials***

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

414 ***Procedures***

415 Study 2 involved participants completing the AOMI-AQ for a second time, within 7-10 days of
416 the first testing session (M duration = 8.02 days, $SD = 2.3$). This 7-10 day period was similar to that used
417 in recent test-retest reliability protocols for other imagery ability measures (Plakoutsis et al., 2023; Suárez
418 Rozo et al., 2022). As shown in Figure 3, for this second session participants were randomly assigned to
419 complete the AOMI-AQ with supervision from a researcher ($n=82$) or without supervision ($n=92$).
420 Supervised testing sessions followed a similar format to Study 1, in which a researcher led the testing
421 session and ensured that participants physically performed the movements instructed in the AOMI-AQ
422 when required. For unsupervised testing sessions, participants were sent a link to the AOMI-AQ by email
423 and instructed to complete the questionnaire independently. This aspect allowed confirmation of whether
424 the self-administration of the questionnaire would be reliable and appropriate for future use. To encourage
425 adherence to the instructions, the self-administered version of the AOMI-AQ included a fixed response

426 delay of 10 seconds during physical performance and AOMI components of the questionnaire to ensure
427 that participants could not continue until sufficient time to perform the actions when required had elapsed.
428 Response times when providing AOMI ratings were also recorded to ensure the video was observed in
429 full. Based on reaction times we checked to ensure adequate time was taken to have performed or
430 imagined the task.

431

432

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

437

438 *Data analysis*

439 Confirmatory factor analysis was used to verify the factor solution from Study 1. The weighted
440 least square mean and variance adjusted estimation method calculated model fit and parameter estimates.
441 This is suitable for data that contains ordinal characteristics (Li, 2016). Fit indices of chi-square,
442 Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Standardized Root-Mean-Square Residual
443 (SRMR) were included when evaluating data-model fit. Good fit thresholds were $CFI \geq 0.95$, $TLI \geq 0.95$,
444 and $SRMR \leq 0.08$ (Hu & Bentler, 1995). Alongside chi-square, the normed chi-square was presented
445 given the sensitivity of chi-square to sample size, with values <1 suggestive of overfitting (Obst & White,
446 2005).

447 To establish the test-retest reliability of the AOMI-AQ we conducted a Pearson's correlation
448 between participants' initial completion of the AOMI-AQ and the retest dataset. In addition, an
449 independent-samples t-test was conducted comparing the scores of participants who completed the retest
450 with the supervision of a researcher and without supervision. We also analyzed potential differences

451 between participants' initial supervised AOMI-AQ responses with their supervised or unsupervised
 452 retests using paired-samples t-tests.

453 **Results**

454 Mean values for the completions of the supervised and unsupervised AOMI-AQ are presented in
 455 Table 2. Confirmatory factor analysis revealed good fit for the hypothesised unidimensional model, χ^2
 456 (20) = 73.21, $p < 0.001$, normed $\chi^2 = 3.66$, CFI = 0.99, TLI = 0.98, SRMR = 0.02. Scrutiny of factor
 457 loadings indicated that all items loaded above 0.6, thus meeting the strict requirements of Hair et al.
 458 (1998). Item loadings ranged from 0.75 to 0.93, with an average loading of 0.85. This supported the
 459 existence of a single dimension underpinning the measure, and findings aligned with the emergence of a
 460 single dimension within Study 1.

461
 462 **Table 2.** Means (M) and standard deviation (SD) for responses for the test-retest of the AOMI-AQ when
 463 supervised and unsupervised.

| Test | Supervision | <i>N</i> | <i>M</i> | <i>SD</i> |
|----------------|--------------|----------|----------|-----------|
| 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 |

464
 465 Cronbach's Alpha for the retest dataset indicated an excellent internal reliability for the AOMI-
 466 AQ ($\alpha = 0.941$), which was similar to the Study 1 dataset ($\alpha = 0.92$). Correlational analysis between these
 467 two datasets showed there to be a significant positive correlation ($r = 0.74, p < 0.001$). Individuals who
 468 rated their AOMI ability highly on the initial test also reported high scores at retest, and the reverse was
 469 true for those who rated lower (see Figure 4).

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

472 = 0.65, $p < 0.001$). An independent samples t-test comparing retest responses for individuals who were
473 supervised ($M = 44.94$, $SD = 8.27$) to those who were unsupervised ($M = 44.95$, $SD = 8.68$) indicated no
474 differences in responses, $t(167.35) = 0.248$, $p = 0.805$. Furthermore, paired-samples t-tests revealed no
475 statistical differences between participants' initial AOMI-AQ responses and their retest response
476 depending on whether they were supervised, $t(81) = 0.131$, $p = 0.896$, or unsupervised, $t(169.78) = 0.312$,
477 $p = 0.756$.

478

479

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

483

484 **Discussion**

485 The aim of Study 2 was twofold: (i) to establish the test-retest reliability of the AOMI-AQ, and
486 (ii) to establish whether differences in AOMI-AQ scores exist when completed under supervision from a
487 researcher or independently by participants. Confirmatory factor analysis confirmed good factor loadings
488 for each item, and in agreement with Study 1, confirmed only one underlying factor. The test-retest
489 reliability findings demonstrate that the AOMI-AQ, when combining supervised and unsupervised retests,
490 has acceptable-to-good test-retest reliability (Cicchetti, 1994; Nunnally, 1978), indicated by a significant
491 positive correlation between test-retest datasets ($r = 0.74$). Isolating these responses to only individuals
492 who were supervised for both testing sessions showed excellent test-retest reliability ($r = 0.83$; Cicchetti,
493 1994; Nunnally, 1978). Evidence that the AOMI-AQ has good-to-excellent test-retest reliability, in which
494 participants' scores on the questionnaire remain stable across testing sessions, enhances the possible
495 applications of the AOMI-AQ in research and applied settings. Specifically, this finding opens up the
496 possibility to use the AOMI-AQ as an outcome measure for researchers or applied practitioners seeking to
497 establish techniques to improve AOMI abilities of participants or clients.

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

506 **General discussion**

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

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

522 Prior to the development of the AOMI-AQ, in the absence of an appropriate measure, researchers
523 investigating AOMI have previously attempted to quantify AOMI ability using pre-existing MI ability

524 questionnaires. Whilst these undoubtedly measure processes related to AOMI, it is recognised the ability
525 to generate MI involves different processes to those required for AOMI (Scott et al., 2022). A major
526 distinction between conventional uses of MI and AOMI is the presence of a visual display during AOMI.
527 This display provides visual content which the individual then uses as a scaffold and stimulus to generate
528 their kinesthetic representation of the observed action (Eaves et al., 2022), which could also be considered
529 an explicit form of what has been referred to as ‘cross-modal imagery’ (Nanay, 2018; Spence & Deroy,
530 2012). While the validation of the AOMI-AQ supports its likeness to the MIQ-3 and VMIQ-2, the
531 discovery of only moderate-to-large positive correlations between the AOMI-AQ and MIQ-3 ($r = 0.71$)
532 and VMIQ-2 ($r = 0.43$) may provide indirect support that different processes were measured, providing
533 tangential justification for a specific measure of AOMI ability.

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

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

550 Eaves, 2018; Dror & Kosslyn, 1994). In Study 1, participants were seemingly unable to distinguish
551 between the concepts of image generation and image maintenance. As this questionnaire was developed
552 based on the MIQ-3, which measures imagery generation abilities, we propose that the AOMI-AQ was
553 accurate in the measurement of this aspect of imagery. While it could be the case that AOMI is not as
554 multidimensional as VI or MI (e.g., Cumming & Eaves, 2018; Dror & Kosslyn, 1994; Kraeutner et al.,
555 2020), and requirements for VI are indeed limited during AOMI (Wright et al., 2022), it may also be that
556 subjective based measures such as self-report questionnaires are not sensitive to the maintenance property
557 of AOMI. Therefore, this aspect of AOMI may best be captured and quantified through other methods
558 such as neurophysiological or chronometry-based measures.

559 The AOMI-AQ provides the first valid and reliable tool by which AOMI ability can be quantified
560 and this has multiple benefits for both research and applied practice. For example, in research contexts,
561 the ability to accurately measure AOMI ability now provides researchers with a tool to (i) introduce
562 participants to the concept of AOMI, (ii) screen for AOMI ability as part of study inclusion/exclusion
563 criteria, (iii) control for AOMI ability or split participants based on AOMI ability when allocating
564 experimental groups, (iv) monitor changes in AOMI ability as an outcome measure in research, and (v)
565 report participant AOMI abilities (Moreno-Verdú et al., 2024), without having to rely on MI ability
566 measures as a proxy measure for AOMI. Similarly, in applied contexts, practitioners can now assess and
567 monitor AOMI ability prior to and throughout AOMI training programmes and interventions
568 administered to their clients. Moreover, in applied settings, the AOMI-AQ could be used alongside
569 previously established MI and AO ability measures to help determine which simulation approach may be
570 easier, more engaging, or better suited to their client, allowing them to tailor simulation interventions
571 based on their clients needs and preferences. In all these cases researchers and practitioners may find it
572 helpful to consider cut-off values for distinguishing ‘good’ and ‘poor’ AOMI ability. While future
573 research may help establish such values, appropriate initial criteria based on the recommendations of
574 Robin and Coudevylle (2018) and Robin and Blandin (2021) could be to consider mean AOMI-AQ scores
575 > 5 and < 2 , respectively, as indicative of good and poor AOMI ability.

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

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

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

611 **References**

- 612 Binks, J. A., Emerson, J. R., Scott, M. W., Wilson, C., Van Schaik, P., & Eaves, D. L. (2023). Enhancing
613 upper-limb neurorehabilitation in chronic stroke survivors using combined action observation and
614 motor imagery therapy. *Frontiers in Neurology*, *14*, 1097422.
615 <https://doi.org/10.3389/fneur.2023.1097422>
- 616 Chye, S., Valappil, A. C., Wright, D. J., Frank, C., Shearer, D. A., Tyler, C. J., ... & Bruton, A. M. (2022).
617 The effects of combined action observation and motor imagery on corticospinal excitability and
618 movement outcomes: Two meta-analyses. *Neuroscience & Biobehavioral Reviews*, *143*, 104911.
619 <https://doi.org/10.1016/j.neubiorev.2022.104911>
- 620 Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized
621 assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290.
622 <https://doi.org/10.1037/1040-3590.6.4.284>
- 623 Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.
- 624 Cumming, J., & Eaves, D. L. (2018). The nature, measurement, and development of imagery ability.
625 *Imagination, Cognition and Personality*, *37*(4), 375–393.
626 <https://doi.org/10.1177/0276236617752439>
- 627 Dance, C. J., Ward, J., & Simner, J. (2021). What is the link between mental imagery and sensory
628 sensitivity? Insights from aphantasia. *Perception*, *50*(9), 757–782.
629 <https://doi.org/10.1177/03010066211042186>
- 630 De Vries, S., Tepper, M., Feenstra, W., Oosterveld, H., Boonstra, A. M., & Otten, B. (2013). Motor imagery
631 ability in stroke patients: the relationship between implicit and explicit motor imagery measures.
632 *Frontiers in human neuroscience*, *7*, 790. <https://doi.org/10.3389/fnhum.2013.00790>
- 633 Dilek, B., Ayhan, C., & Yakut, Y. (2020). Reliability and validity of the Turkish version of the movement
634 imagery questionnaire-3: Its cultural adaptation and psychometric properties. *Neurological
635 Sciences and Neurophysiology*, *37*(4), 221. https://doi.org/10.4103/NSN.NSN_30_20

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

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

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

644 Eaves, D. L., Hodges, N. J., Buckingham, G., Buccino, G., & Vogt, S. (2022). Enhancing motor imagery
645 practice using synchronous action observation. *Psychological Research*, 1-17.
646 <https://doi.org/10.1007/s00426-022-01768-7>

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

693 Martini, R., Carter, M. J., Yoxon, E., Cumming, J., & Ste-Marie, D. M. (2016). Development and validation
694 of the Movement Imagery Questionnaire for Children (MIQ-C). *Psychology of Sport and Exercise*,
695 22, 190-201. <https://doi.org/10.1016/j.psychsport.2015.08.008>

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

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

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

705 Moreno-Verdú, M., Hamoline, G., Van Caenegem, E. E., Waltzing, B. M., Forest, S., Valappil, A. C., ... &
706 Hardwick, R. M. (2023). Guidelines for Reporting Action Simulation Studies (GRASS): proposals
707 to improve reporting of research in Motor Imagery and Action Observation. *Neuropsychologia*,
708 108733. <https://doi.org/10.1016/j.neuropsychologia.2023.108733>

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

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

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

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

716 Olsson, C. J., & Nyberg, L. (2010). Motor imagery: if you can't do it, you won't think it. *Scandinavian*
717 *journal of medicine & science in sports*, 20(5), 711-715. <https://doi.org/10.1111/j.1600->
718 0838.2010.01101.x

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

721 O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel
722 analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, and Computers*, 32,
723 396–402. <https://doi.org/10.3758/BF03200807>

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

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

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

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

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

- 738 Roberts, R., Callow, N., Hardy, L., Markland, D., & Bringer, J. (2008). Movement imagery ability:
739 development and assessment of a revised version of the vividness of movement imagery
740 questionnaire. *Journal of Sport and Exercise Psychology*, 30(2), 200–221.
741 <https://doi.org/10.1123/jsep.30.2.200>
- 742 Robin, N., & Blandin, Y. (2021). Imagery ability classification: Commentary on «kinaesthetic imagery
743 ability moderates the effect of an AO+MI intervention on golf putt performance: A pilot study» by
744 McNeill, Ramsbottom, Toth, and Campbell (2020). *Psychology of Sport and Exercise*, 57, 102030.
745 <https://doi.org/10.1016/j.psychsport.2021.102030>
- 746 Robin, N., & Coudevylle, G. R. (2018). The influences of tropical climate, imagery ability, distance and
747 load on walking time. *International Journal of Sport Psychology*, 3, 66-87.
748 <https://doi.org/10.1080/20445911.2017.1384382>
- 749 Romano-Smith, S., Roberts, J. W., Wood, G., Coyles, G., & Wakefield, C. J. (2022). Simultaneous and
750 alternate combinations of action-observation and motor imagery involve a common lower-level
751 sensorimotor process. *Psychology of Sport and Exercise*, 63, 102275.
752 <https://doi.org/10.1016/j.psychsport.2022.102275>
- 753 Rungsirasilp, N., & Wongsawat, Y. (2022). Applying Combined Action Observation and Motor Imagery to
754 Enhance Classification Performance in a Brain–Computer Interface System for Stroke Patients.
755 *IEEE Access*, 10, 73145-73155. <https://doi.org/10.1109/ACCESS.2022.3190798>
- 756 Scott, M., Taylor, S., Chesterton, P., Vogt, S., & Eaves, D. L. (2018). Motor imagery during action
757 observation increases eccentric hamstring force: an acute non-physical intervention. *Disability and
758 Rehabilitation*, 40(12), 1443–1451. <https://doi.org/10.1016/j.ajsep.2022.07.002>
- 759 Scott, M. W., Wright, D. J., Smith, D., & Holmes, P. S. (2022). Twenty years of PETTLEP imagery: An
760 update and new direction for simulation-based training. *Asian Journal of Sport and Exercise
761 Psychology*, 2(2), 70–79. <https://doi.org/10.1016/j.ajsep.2022.07.002>
- 762 Scott, M. W., Wood, G., Holmes, P. S., Marshall, B., Williams, J., & Wright, D. J. (2023). Combined action
763 observation and motor imagery improves learning of activities of daily living in children with

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

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

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

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

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

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

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

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

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

789 3 (MIQ-3). *Musculoskeletal Science and Practice*, 51, 102313.
790 <https://doi.org/10.1016/j.msksp.2020.102313>

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

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

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

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

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

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

811 Ziv, G., Lidor, R., Arnon, M., & Zeev, A. (2017). The Vividness of Movement Imagery Questionnaire
812 (VMIQ-2)–translation and reliability of a Hebrew version. *Israeli Journal of Psychiatry*, 54, 48-
813 53.