1   **The Comparative Suitability of Traditional and Task-specific Think Aloud Training**

2

3

4   Phil D. J. Birch1 & Amy E. Whitehead2

5   1Institute of Sport, University of Chichester, Chichester, UK

6   2School of Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK

7

8   Corresponding author: Phil D. J. Birch, University of Chichester, College Lane, Chichester, West Sussex, PO19

9   6PE, UK. Email: P.Birch@chi.ac.uk.

10

11

15 **Abstract**

16 The Think Aloud (TA) protocol is used to capture conscious cognition for wide ranging applications. However,

17 the methods used to train the TA technique have been inconsistent involving a mixture of both traditional

18 guidelines (Ericsson & Simon, 1993) and task-specific examples. This study aimed to examine how best to train

19 the TA process. We recruited 20 competitive golfers as research participants, and we randomly assigned them

20 to equal sized groups of traditional TA training as described by Ericsson and Simon (1993) and task-specific

21 training in which participants were familiarized with TA via task-specific examples. Following training, all

22 participants performed a golf task and were asked to TA. We transcribed audiotapes of their verbatim TA

23 content and analyzed them using a deductive framework. We also collected various social validation self-report

24 measures to assess participant perceptions of TA training. Overall, we found no significant differences in the

25 frequency or type of TA verbalizations when comparing traditional and task-specific TA training groups.

26 However, participants in the task-specific training group reported more favourable perceptions of training and

27 found training significantly clearer than did participants in the traditional training group. We suggest that these

28 findings support traditional TA training following Ericsson and Simon's (1993) training guidelines, but adding

29 task-specific examples seems to increase the familiarity of TA use and facilitate more reliable and accurate

30 cognition data for research use.

31

32 Key words: Think aloud protocol, cognition, training, golf.

33

34

35   **Introduction**

36   Verbal reports from the Think Aloud (TA) protocol (Ericsson & Simon, 1993) have been used for

37   decades to capture problem solving and decision-making data (e.g., Bloom & Broder, 1950) in a variety of

38   applications. TA has been used in medical settings in relation to pain management (e.g., Taylor, Allsop, Bewick,

39   & Bennett, 2016), surgery (e.g., McRobert et al., 2013), nursing (e.g., Banning, 2008), teaching (e.g., Ellis,

40   2013) and within various sports to capture in-performance cognitions (e.g., Ward, Williams, & Ericsson, 2003;

41   Whitehead, Taylor, & Polman, 2016). The terminology surrounding both TA and its verbal reports have been

42   used interchangeably, with some researchers preferring the term TA (e.g., Welsh, Dewhurst, & Perry, 2018) and

43   others using terms related to the specific time the reports were gathered, such as immediate and retrospective

44   verbal reports (e.g., McRobert et al., 2013). While TA is a verbal report method that captures in-event

45   cognitions, within this paper we use the term TA as an umbrella term for discussing past literature, and we offer

46   greater detail regarding our use of the TA method.

47   TA requires a performer to verbalize his or her thought process continuously while performing.

48   Ericsson and Simon (1993) proposed three TA levels: (a) Level 1: simple vocalisations of inner speech in which

49   the individual makes no effort to communicate his or her thoughts; (b) Level 2: verbal encoding and vocalization

50   of an internal mental representation that was not originally verbally coded (e.g., verbal encoding and vocalizing

51   scents, visual stimuli, or movements) and conveys only the information that is in the participant's focus; and (c)

52   Level 3: explanations of the individual's thoughts, ideas, hypotheses, or motives (Ericsson & Simon, 1993). For

53   example, in Level 3 verbalizations, the performer might explain why a certain medical procedure should be

54   conducted or why a certain golf shot or club was selected. It is important to note that most TA researchers have

55   opted to study Level 2 TA verbalizations because Level 2 verbalizations capture an individual's ongoing

56   cognition within his/her short-term memory (STM) and are not obtained retrospectively from long-term memory

57   (LTM). Retrieving information from LTM may slow the TA process and make the obtained data less naturally

58   procured. During the performance of an activity, information in STM is only briefly available (Newell & Simon,

59   1972). As the task continues and new information is presented, previous information is lost, affording the

60   individual an opportunity to retrieve and verbalize information that is not directly needed for task performance

61   and yields data that is not a product of any cognitive process that mediates the performance (Eccles, 2012).

62   The benefit of using TA is that it allows data regarding thought processes to be captured within real

63   time, reducing risks of memory decay and retrospective bias (e.g., Folkman & Moskowitz, 2004; Ptacek, Smith,

64   Espe, & Raffety, 1994; Smith, Leffingwell, & Ptacek, 1999; Stone et al., 1998). Furthermore, TA allows

65 researchers to identify potential differences in perceptual-cognitive processes between performers of varying

66 expertise (Williams, Ford, Eccles, & Ward, 2011). For example, in early work, DeGroot (1978) demonstrated

67 TA evidence of domain specific knowledge and mental representation by asking master and intermediate chess

68 players to reconstruct the locations of chess pieces after viewing the board for only a few seconds. While

69 master chess players showed significant superiority at this skill, compared to intermediate players, there were no

70 group differences when chess piece locations were presented randomly on the board. This experiment showed

71 that, over time and practice, the higher skilled chess players had stored thousands of chunks of chess related

72 information (a chunk was defined as a sequence of pieces with between piece intervals of less than two seconds)

73 and could retrieve this information from LTM to give them greater familiarity with and easier recall of

74 previously seen chess patterns after only a few seconds of viewing. However, they lost this advantage when

75 random distributions of pieces were unrelated to this knowledge base.

76   Research using TA has identified cognitive differences between various levels of performers in a wide

77 range of domains. Within medical research, McRobert et al. (2013) found that skilled physicians demonstrated

78 higher diagnostic accuracy and selected better quality options during diagnostic reasoning, compared to less

79 skilled physicians. In chess, researchers identified that Grandmaster players search more quickly and have

80 superior pattern recognition compared to lower level players (Connors, Burns, & Compitelli, 2011).

81 Furthermore, Whitehead, Taylor and Polman (2015) found that thought processes of lower skilled golfers

82 reflected greater focus on technical performance mechanics, whereas higher skilled golfers focused more on

83 execution planning. Information gleaned through TA research methods may enable various practitioners to

84 identify potential flaws in their cognitive strategies, and TA training can be an effective intervention for

85 enhancing performance.

86   Despite Ericsson and Simon's (1993) emphasis on the importance of TA training, their general

87 instructions included one simple mental arithmetic task and one problem-solving task to familiarize participants

88 with the TA technique. Specifically, Ericsson and Simon (1993, p. 375-379) stated:

89 "Good, before we turn to the real experiment, we will start with a couple of practice problems. I want you to talk

90 aloud while you do these problems. First, I will ask you to multiply two numbers in your head. So *talk aloud*

91 while you multiply 24 times 34. Good! Now I would like you to solve an anagram. I will show you a card with

92 scrambled letters. It is your task to find an English word that consists of all the presented letters. For example, if

93 the scrambled letters are KORO, you may see that these letters spell the word ROOK. Any questions? Please

94 "talk aloud" while you solve the following anagram! <NPEPHA = HAPPEN>."

95   Adaptations of these warm-up tasks (see Eccles, 2012; Ericsson & Kirk, 2001) have also been used in many

96   other studies that utilized the TA procedure (e.g., Aitken & Mardegan, 2000; Nicholls & Polman, 2008).

97   However, these tasks are not task specific, and it is unknown to what extent participants believe that these tasks

98   fully equip them with the ability and confidence to effectively perform TA. Indeed, Van Someren, Barnard and

99   Sanberg (1994) highlighted the importance of aligning the training task to the target task, or as they state, "...in

100   general it is wise to look for a task which is not too different from the target task" (p. 43).

101   When learning a new skill, domain specificity is extremely important, especially in information

102   processing. When a new skill is being processed, the body (one or more of the sense organs) identifies the

103   task/stimuli and a response is selected, prepared and initiated. This process involves internal memorialized

104   representations (De Groot, 1978; Elliot et al., 2010). During this new activity/engagement with a new stimulus,

105   the activity is coded within the brain and identified as new or familiar, according to its similarity to other mental

106   representations already stored in LTM. Lord and Maher (1991) provided a simplistic explanation for

107   information processing associated with how a task is performed. Their view emphasized the energy required to

108   perform the task. More specifically, the number of tasks that can be performed concurrently is limited by the

109   combined amount of energy that tasks consume (Anderson 1990; Kahneman, 1973). The energy requirements

110   needed to perform a task depend on how well the task has been practiced. Therefore, novel tasks require much

111   more energy or attention (controlled processing) while well-rehearsed tasks require fewer attentional demands

112   (automatic processing). It could be argued that if a task, such as learning TA, is closely linked to the

113   performance domain, then the energy to perform the task within this familiar environment may be less than if

114   the task is not domain specific. Intuitively then, when learning TA in a specific environment, we might predict

115   that the learning process will be easier with task-specific examples that allow connections to be made with task-

116   specific representations already stored in LTM.

117   In an effort to supplement the traditional TA training methods recommended by Ericsson and Simon

118   (1993, p. 375-379), some researchers have added task-specific warm-up tasks to better familiarize participants

119   with TA. North, Ward, Ericsson and Williams (2011) provided the following information; "...several domain-

120   specific examples were included as part of the training protocol. The training session included instruction and

121   practice at thinking aloud, and retrospectively reporting these thoughts using a range of generic problems and

122   task-specific video-based scenarios" (p. 160-161). In Arsal, Eccles and Ericsson's (2016) study, participants

123   "...practiced thinking aloud while putting twice over 89 cm" (p. 21). Runswick, Roca, Williams and Bezodis

124   (2018) stated, "...training included instruction on thinking aloud and giving immediate retrospective verbal

125   reports by solving a range of generic and domain-specific tasks" (p. 711). Similarly, Calmeiro and Tenenbaum's

126   (2011, p 226) second phase of TA training "...consisted of verbalization practice while putting" and Whitehead

127   et al.'s (2015, p. 3-4) TA protocols were "...adapted to golf putting based upon the guidelines set out by Ericsson

128   and Simon (1993) and Nicholls and Polman (2008)." Despite Whitehead et al. (2017:18) providing participants

129   with task-specific (cycling) video material prior to data collection, it is not entirely clear what this involved.

130   Although it is positive to see that some task-specific TA training has been implemented in past studies, TA

131   training procedures may be further strengthened by more consistent use of task-specific warm-up tasks to ensure

132   that TA training can be replicated in follow-up research. Enhancing the specificity of TA instructions could lead

133   to a number of favourable outcomes. Firstly, as noted, more specific instructions might increase other

134   researchers' understanding of how TA was trained, and thus, enable its replication. To date, the literature affords

135   limited information for follow-up researchers. Indeed, Samson, Simpson, Kamphoff and Langlier (2015, p. 11)

136   conceded that a limitation of their study was the "...non-sport nature of the warm-up tasks," and they encouraged

137   researchers to examine the effectiveness of TA training protocols. Secondly, greater TA instruction specificity

138   might increase the participant's ability to learn and use TA effectively, possibly enhancing the quality of

139   verbalizations captured. Given the importance placed on upholding data gathering rigour in TA research

140   (Ericsson & Simon, 1993), greater understanding is needed concerning the precise procedures utilized to train

141   TA. These circumstances suggest a need for further research to examine optimal TA training methods and their

142   impact on athlete verbalizations.

143   To shed further light on TA training effectiveness, it would appear intuitive to ask athletes for their

144   opinions regarding the training process. To the author's knowledge, previous research has not explicitly

145   examined how participants perceive TA training. Traditionally, social validation procedures have been used to

146   measure participant perceptions and satisfaction with an intervention (e.g., Mellalieu, Hanton, & O'Brien, 2009;

147   Thelwell, Greenlees, & Weston, 2006). Consequently, social validation affords a method for examining athlete

148   perceptions of the respective components of TA training (e.g., clarity of verbal instructions, effectiveness of

149   training exercises), and in turn, the effectiveness of TA training methods. Further research to examine methods

150   of training TA may afford a more consistent approach to using TA, perhaps leading to a more in depth

151   understanding of expert performers' cognitive processes across domains. Due to the exploratory nature of this

152   paper and a dearth of explicit investigations of how TA is trained, we aimed to examine the impact of traditional

153   and task-specific TA training procedures on both participants' cognitive processes and their perceptions of

154   training effectiveness. Given that more positive perceptions of TA training (e.g., confidence of using TA) might

155   be associated with a higher willingness to verbalize one's cognitions, we hypothesised that task-specific TA

156   training would result in significantly more verbalizations than would traditional TA training. Given that task-

157   specific training may promote greater storage of contextual information in the LTM, we also hypothesized that

158   task-specific training would result in more favourable perceptions of TA training effectiveness compared to

159   traditional TA training.

160   **Method**

161   **Participants**

162   We recruited 20 golfers from a golf club in the South of England and split them into two equal sized

163   groups with comparable skills. We then randomly assigned the groups to either traditional TA training ($n = 10$;

164   six males, four females; age: $M = 42.7$ years, $SD = 11.8$; golf handicap: $M = 13.1$, $SD = 10.4$) or task-specific

165   TA training ($n = 10$; 10 males, zero females; age: $M = 43.0$ years, $SD = 14.2$; handicap: $M = 12.5$, $SD = 10.3$).

166   Participants in the traditional TA training group had an average of 11.2 ($SD = 9.6$) years of competitive playing

167   experience, played at least once per week and had played in an average of 19.6 ($SD = 11.8$) competitions in the

168   12 months leading up to their study participation. Participants in the task-specific TA training group had an

169   average of 19.7 ($SD = 9.0$) years of competitive playing experience, played at least once per week and had

170   played an average of 11.1 ($SD = 7.1$) competitions in the 12 months leading up to their participation in the

171   study. No participants had TA experience prior to participating within this study. All participants identified their

172   ethnicity as white British. We secured institutional ethical approval for the study protocol, and we obtained

173   informed consent from all participants prior to their participation.

174   **Materials**

175   **TA training videos.** All participants used their own golf clubs and balls to perform the golf task,

176   conducted on a practice green at their home golf course. We used a Sony HXR-NX30N camcorder with radio

177   microphone to record participant verbalizations. The mini radio microphone was attached to the participant's

178   collar, and we placed a wire inside the shirt connecting to the recording device placed in the participants pocket.

179   The stimuli used in this experiment were two TA training videos, each consisting of visual and verbal

180   instructions on how to perform TA (see Table 1 for content summary of each video). The purpose of the videos

181   was to provide participants with an understanding of how TA works so that they could competently perform the

182   technique. In line with Ericsson and Simon's (1993) guidelines, both videos provided identical instructions to

183   train participants in performing TA. Example instructions included, "Think aloud involves you saying out loud

184   everything that you are thinking as you are performing the task," and "It is important that you think aloud all

185   your thoughts as best as you can during that time." Given that this study aimed to examine Level 2 TA, both

186   training videos included instructions to promote Level 2 TA and deter Level 3 TA. In accordance with Ericsson

187   and Simon's (1993) guidelines, the videos stated, "I don't want you to try to plan out what you say or try to

188   explain to me what you are saying." In order to promote authentic projection of thoughts (Ericsson & Simon,

189   1993), both videos instructed participants to, "Just act as if you are alone speaking to yourself." To ensure that

190   participants were performing TA throughout the golf task (Ericsson & Simon, 1993), the videos stated, "It is

191   most important that you keep talking. If you are silent for any long period of time, I will ask you to talk."

192   Further instructions were also written specifically for this study. These included, "We are interested in knowing

193   your thoughts as they come to mind during the golf task. This includes the thoughts you have in the lead up to

194   hitting the ball, whilst the ball is in motion, after the ball has come to rest, and as you walk to play your next

195   shot," and "Everything you say is confidential - the researcher will not judge your thoughts and please use swear

196   words if you feel necessary." It is important to note that participants were instructed to refrain from verbalizing

197   during skill execution to reduce possible interference with motor movement (Schmidt & Wrisberg, 2000).

198   The remainder of the videos consisted of the participants' TA training treatment: traditional TA

199   training or task-specific TA training. The training exercises used in the traditional TA training video were based

200   on the recommendations of Ericsson and Simon (1993) and consisted of three different groups of tasks: (a) four

201   alphabetical problem-solving tasks (e.g., what is the fourth letter after H); (b) five tasks counting the number of

202   dots on a page; and (c) two general problem-solving tasks (e.g., name two vegetables that begin with the letter

203   C). These training exercises have been used in a number of previous research studies (e.g., Nicholls & Polman,

204   2008; Samson et al., 2015; Whitehead et al., 2015:16).

205   The exercises used in the task-specific training video were developed for this study and consisted of

206   three different golf scenarios: (a) Tee shot on a par 5 hole; (b) fairway (second) shot on a par 5 hole; (c)

207   greenside approach (third) shot over a bunker. For the first scenario, we provided the following information: "It

208   is the first hole of a monthly medal. You are standing on the first tee of a 473 yard par 5. You have been striking

209   the ball very well and scoring very well in the lead up to this competition. It is a reasonably warm summer's day

210   and the course is firm and playing fast. The weather is overcast and there is a strong wind against." For the

211   second scenario, the following information was provided: "You are now playing your second shot on the same

212   hole in the monthly medal. Again, you have been striking the ball very well and scoring very well in the lead up

213   to this competition. It is a reasonably warm summer's day and the course is firm and playing fast. The weather

214   is overcast and there is a strong wind against. The pin is cut back right. Your ball is in the middle of the fairway

215   and lying very nicely. Your ball is marked by the white ball on the right diagram." For the third scenario, the

216   following information was provided: "You are now playing your third shot on the same hole in the monthly

217   medal. Your short game has been poor in the lead up to this competition. It is a reasonably warm summer's day

218   and the greens are playing firm and fast. The weather is overcast and there is a strong wind against. The flag is

219   cut back right. Your ball is lying poorly in the left rough – marked by the white ball on the right diagram."

220   Previous research has incorporated similar task-specific TA training exercises (e.g., Calmeiro & Tenenbaum,

221   2011; North et al., 2011; Runswick et al., 2018).

222   At the end of each description for the respective scenarios, participants were instructed, "Please use the

223   information in the diagrams and tell us your thoughts on how you would play this shot." At this moment, two

224   diagrams appeared on the video to help facilitate TA. The diagram on the left provided a bird's eye view of the

225   hole and the yardages to and from its respective features (e.g., yardage to the bunker from the tee). The diagram

226   on the right was a first-person view of the hole (albeit from an elevated position) and represented the

227   information a golfer would gain whilst performing on a golf course. Once the participants received their TA

228   training treatment, both groups were instructed to complete the TA training checklist to assess how well the

229   participant had learned the requirements of TA. Finally, all participants were instructed to have three practice

230   trials on the golf task while verbalizing to familiarize themselves thinking aloud. The traditional TA training

231   video was 4:47 minutes in duration and the task-specific TA training video was 4:14 minutes in duration.

232   [Insert Table 1 about here.]

233   **The golf task**. The golf task was specifically designed for this study as a means to facilitate authentic

234   short game golf shots (i.e., chipping and putting) that golfers would typically face during a round of golf. Given

235   that every shot is different whilst playing a round of golf, we used three different hitting zones (see Figure 1).

236   Hitting zone 1 was positioned on an up-hill lie 15 meters from the hole and exhibited an incline. Hitting zone 2

237   was positioned on a flat lie 19 meters from the hole and exhibited an incline. Hitting zone 3 was positioned on a

238   side-hill lie (ball below participant's feet for a right-handed golfer) 11 meters from the hole and exhibited a

239   decline. All hitting zones were located on shortly mown grass and participants were permitted to place their ball

240   within a one-meter squared area. The speed of the green was measured on a Stimpmeter. The total amount of

241   feet the ball rolls from the Stimpmeter gives an approximation as to the pace of the green. The green measured

242   an average of nine on the Stimpmeter. Participants were required to hit the ball in the hole in as few shots as

243   possible and were allowed to select which club to use. In order to enhance the ecological validity of the task, a

244   series of pressure manipulations were enforced (Baumeister & Showers, 1986). Participants were informed that

245   they would be entered into a competition in which the participant with lowest score would receive three

246   premium golf balls. Participants were also informed that their performance scores (i.e., amount of shots taken)

247   would be published on a leader board that would be readily available for other participants to see before they

248   performed their trials. Indeed, participants were informed of other participants' scores before performing to

249   facilitate the comparative and evaluative nature of the task.

250   [Insert Figure 1 about here.]

251   **Measures**

252   **TA protocol.** We recorded Level 2 verbalizations during the golf task. Participant verbalizations were

253   transcribed verbatim.

254   **Task commitment.** We measured task commitment to determine the level of engagement with the task

255   and to determine if there were differences in task engagement between participants in the two TA training

256   groups. In accordance with research by Arsal (2013), we used the following item: "How committed were you to

257   the task while performing?" Participants were instructed to rate their commitment on a scale, with 10%

258   increments ranging from 0% (*not at all*) to 100% (*very much*).

259   **TA training checklist.** We designed the TA training checklist specifically for this study and required

260   participants to recall seven key training components to successfully perform TA: (a) all verbalizations were

261   confidential; (b) all thoughts were to be spoken; (c) refrain from explaining your thoughts; (d) use TA before

262   and after your shot; (e) refrain from verbalizing during skill execution; (f) periods of silence will result in being

263   prompted; and (g) swearing is permitted. This was used to assess how well the participants had learned the

264   requirements of TA.

265   **Self-efficacy.** Self-efficacy for thinking aloud was measured to determine participants' beliefs in

266   thinking aloud whilst performing the golf task. In accordance with Bandura's (1997) recommendation,

267   participants indicated the strength of their self-efficacy for thinking aloud concurrently to performing the golf

268   task by responding on a one-item Likert-type scale with 10% increments ranging from 0% (*not at all confident*)

269   to 100% (*completely confident*).

270   **TA social validation.** Social validation procedures have been suggested as a means of strengthening

271   the external validity of technical and practical action research by offering personal insight into the intervention

272   through the participants' experiences (Newton & Burgess, 2008). Social validity refers to the "consideration of

273   social criteria for evaluating the focus of treatment, the procedures that are used and the effects that they have"

274   (Kazdin, 1982, p.479). Furthermore, social validation has been defined as a "supplemental method that

275   facilitates involvement of multiple participants in the evaluation process" (Busse, Kratochwill, & Eilliott, 1995,

276   p. 273). This study has therefore adopted a social validation approach to understand whether the participants

277   considered the TA training procedures to be effective or acceptable (Kazdin, 1982; Wolf, 1978). In accordance

278   with Page and Thelwell's (2013) guidelines, we used quantitative social validation questions in an effort to

279   better understand participants' experiences in using TA. Participants were asked the following questions: (a) Did

280   you enjoy the TA training? (with responses ranging from 1 = *not at all enjoyable* to 7 = *very enjoyable*); (b)

281   How clear were the instructions in the TA training video? (with responses ranging from 1 = *unclear* to 7 = very

282   clear); (c) With regards to helping you learn TA, how effective were the TA practice tasks in the training video?

283   (with responses ranging from 1 = *not at all effective* to 7 = *very effective*); (d) With regards to helping you learn

284   TA, how effective were the physical TA practice trials? (with responses ranging from 1 = *not at all effective* to 7

285   = *very effective*); and (e) Overall, how effective did you think the training was in preparing you to TA during the

286   golf task? (with responses ranging from 1 = *not at all effective* to 7 = *very effective*). Participants were also

287   asked the following qualitatively orientated open questions: (a) Is there anything that you would add to the TA

288   training? And (b) Do you have any further comments regarding the TA training?

289   **Procedure**

290   **Pilot study.** In a pilot study prior to beginning this investigation, we recruited two moderately skilled

291   golfers with handicaps of 7 and 10 and accumulated competitive playing experience of 12 and 10 years,

292   respectively. Both golfers completed the entirety of the experimental procedure, with one receiving the

293   traditional TA training and one receiving the task-specific TA training. Based on their feedback, participants

294   were confident that they could verbalize whilst performing the golf task and that the equipment did not hinder

295   their performance. Participants stated that the golf task was a realistic task which translated well to the golf

296   course.

297   **Experimental procedure.** Prior to conducting the experimental procedure, all participants completed a

298   demographic questionnaire and gave their written informed consent. All participants performed a total of 15

299   practice trials comprising of five trials from the three different hitting zones to familiarize themselves with the

300   demands of the golf task (see Figure 1). Trials were performed sequentially (hitting zone 1, hitting zone 2,

301   hitting zone and so forth) to decrease the likelihood of boredom. We decided from the pilot testing that 15

302   practice trials were appropriate, as this provided sufficient time (without being too laborious) for participants to

303   warm-up and familiarize themselves with the practice green. Each trial on the golf task required the participant

304   to place the ball in the hitting zone, perform their usual pre-performance routine, hit the approach shot as they

305   would on the golf course, walk up to where the ball finished, perform their usual pre-performance putting

306   routine and attempt to putt the ball into the hole in as few shots as possible. Participants were permitted to

307   change their clubs as they normally would.

308   Participants then received their TA training video using an Apple iPad and Sony MDR ZX660AP

309   headphones. Participants were required to complete the TA training checklist. In order to give participants an

310   opportunity to practice using TA whist performing, we gave them three practice trials. During this time, the

311   researcher ensured the participant was competently using TA in line with the instructions given in the training

312   video. Participants were then asked to rate their level of self-efficacy in thinking aloud whilst performing the

313   golf task. Participants completed a series of nine trials on the golf task whilst thinking aloud. Participants were

314   reminded to use TA throughout the nine trials and were told that if they were silent for a period longer than five

315   seconds, they would be asked to resume thinking aloud. Although previous research has used 20 second (e.g.,

316   Nicholls & Polman, 2008) and 10 second (e.g., Whitehead et al., 2015) prompt durations to ensure the

317   occurrence of verbalizations, the pilot study revealed the need to use a shorter duration prompt due to the

318   relatively short gaps between skill executions on the golf task. A researcher walked to the side of the

319   participants (approximately five meters) during the golf task, and there was no communication except that the

320   investigator reminded the participants to continue thinking aloud and what zone to hit from next (Nicholls &

321   Polman, 2008). Other than the presence of the researcher, each participant performed alone. Participants were

322   asked to rate their level of commitment with the golf task after the 15 practice trials and after the nine trials of

323   thinking aloud. At the completion of the think aloud trials, participants completed the social validation questions

324   and the self-efficacy scale (to assess efficacy of using TA in the future).

325   **Data analysis**

326   We analyzed quantitative data gleaned from the task commitment scale, TA training checklist, self-

327   efficacy scale and the social validation questions using SPSS Statistics 23. Given that the data were normally

328   distributed, we conducted a series of independent samples t-tests to examine differences between the traditional

329   TA training group and the task-specific training group.

330   We transcribed TA verbalizations verbatim and subjected them to line by line content analysis. Given

331   the non-anticipatory nature of the golf task used in this present study, we used a golf-specific adapted

332   framework from Calmeiro and Tenenbaum (2011) and Whitehead et al. (2016) to code the verbalizations (see

333   Table 2). The first author analyzed a 10% sample of the data and found an inter-rater reliability agreement of

334   85% (see MacPhail, Koza, & Abler, 2016). Both authors engaged in discussion and came to an agreement for

335   rating discrepancies of the remaining 15%. Since the data were non-normally distributed, we used Mann-

336   Whitney U tests to analyze the differences in themes verbalized between participants in the traditional TA

337   training group and the task-specific training group. We calculated Cohen's (1994) *d* effect sizes to establish the

338   magnitude of differences between the traditionally trained and task-specific trained participants.

339   [Insert Table 2 about here.]

340   The second author independently analyzed the qualitative social validation data to ensure content

341   familiarity. We used inductive content analysis to analyze these data (Scanlan, Stein, & Ravizza, 1989).

342   Following previous research investigations of participants' perceptions of using TA (Whitehead et al., 2018), we

343   employed inductive reasoning to allow verbalization themes to emerge from the data and determined that these

344   data generated three themes. To ensure rigour, the lead author then acted as a critical friend to ensure that data

345   collection and analyses were plausible and defendable (Smith & McGannon, 2017).

346   **Results**

347   **Content of TA data**

348   A comparison of the total verbalization frequency between traditional ($n = 720$, $M = 71.9$, $SD = 20.70$)

349   and task-specific ($n = 719$, $M = 72.00$, $SD = 21.03$) TA training found no significant frequency difference. A

350   series of Mann-Whitney U tests were conducted to investigate the content of the verbalizations of the traditional

351   training group and the task-specific training group (see Table 3 for descriptive statistics) and found no

352   significant verbalization frequency differences when comparing the following verbalized thematic content:

353   gathering information, planning, mental readiness, reactive comments, description of outcome, diagnosis of

354   outcome, and technical information.

355   **Task commitment**

356   Independent samples *t*-tests showed no significant training group difference in post practice

357   commitment check scores or post-TA trial commitment check scores.

358   **TA training checklist**

359   An independent samples *t*-test showed no significant training group difference in the amount of TA

360   instructions recalled.

361   **Self-efficacy**

362   An independent samples *t*-test showed no significant training group difference in either post-practice

363   perceived self-efficacy for using TA or future use perceived self-efficacy in using TA.

364   **Social validation - quantitative**

365   An independent samples *t*-test showed a significant difference, $t(9) = 2.377$, $p = .041$, $d = 1.063$, in

366   perceptions of instruction clarity between the traditional TA training group ($M = 6.10$, $SD = 1.20$) and the task-

367   specific TA training group ($M = 7.00$, $SD = 0$), that favoured the task-specific TA training group. No significant

368   training group differences were found when comparing the remaining social validation questions (see Table 3

369   for descriptive statistics) for enjoyment of using TA scores, effectiveness of the in-video TA training tasks,

370   effectiveness of the physical TA practice trials, or overall TA training effectiveness.

371   [Insert Table 3 about here.]

372   **Social validation - qualitative**

373   Analyses of verbal responses revealed three main themes within these data: Confidence, Task

374   Understanding, and Further Support. Within these themes it was apparent that participants in the traditional

375   training group and the task-specific training group exhibited different qualitative thoughts about their training,

376   as detailed below.

377   **Confidence.** Both the traditional training group and the task-specific training group reported being

378   confident in their use of TA. However, within the traditional training group, some participants reported that they

379   may not have always verbalized everything that they would be thinking, as they were not always comfortable

380   disclosing their thoughts. Participant #9 (traditional TA group) stated, "Some things I did not say, as I was not

381   fully familiar with the task and not used to blurting things out." Conversely, the task-specific group exhibited

382   confidence in their ability to follow the training and use TA. For example, participant #4 (task-specific TA

383   training group) stated, "They were good because they played as a scenario that I could think and apply it to my

384   own ability."

385   **Task understanding.** Both the traditional and the task-specific training groups reported a general

386   consensus that they understood the training tasks they were given. However, within the traditional TA training

387   group some participants reported losing their way and questioned some of the TA training tasks. For example,

388   participant #11 (traditional TA training group) stated, "I lost my way a bit through the training," and participant

389   #13 stated, "The dots were effective, but the other parts of the task, not so much." Participant #10 (task-specific

390   TA training group) reported, "Yeah, it just gets you into the mode of thinking, with a prompt here or there if I

391   wasn't or when I should be doing. So it was good. Very helpful." Participant 14 (task-specific TA training)

392   stated, "I just thought it was pretty simple. It just wasn't too complicated as well, and I was clear about what I

393   had to do."

394   **Future support.** Both TA training groups suggested that supplementary support would aid their ability

395   to use TA proficiently. However, the recommendations provided were different, depending on the type of

396   training received. Participants in the task-specific training group generally reported that they would like to have

397   had more support in using TA in the future and be reassured that the process requires all thoughts to be

398   verbalized. For example, participant #10 (task-specific TA training group) stated, "I'd probably like to do more

399   of it as it was a really good learning tool." Furthermore, participants in the task-specific TA training group also

400   reported that they would like to be reminded more, within the initial training, that all thoughts, no matter how

401   obscure, should be verbalized. For example, participant #2 (task-specific TA training group) stated, "I would

402   have liked to be re-assured more that even strange thoughts should be spoken out loud." Conversely,

403   participants in the traditional TA training group reported that they needed feedback as to whether they were

404   doing TA properly and would have liked more exercises linking TA to the golf environment. For example,

405   participant #11 stated, "I would have liked more comments around the process of TA," and participant 2 (both

406   traditional TA training group) specified, "It could have been more golf related." Furthermore, participant #15

407   (traditional TA training group) expressed the need for a clearer link to golf by stating, "I would have liked more

408   feedback in terms of how the thinking aloud will then relate to golf and if I'm doing it properly."

409   **Discussion**

410   The first aim of this study was to investigate whether the form or type of TA training would impact

411   verbalization frequency. Our results led us to reject our hypothesis that the task-specific TA training would

412   result in significantly more verbalizations when compared to the traditional TA training. We found no

413   significant differences in verbalization frequency across the analyzed categories between the traditional TA

414   training group and the task-specific TA training group (see Table 3). According to information processing

415   theorists (e.g., De Groot, 1978; Elliot et al., 2010), familiarity of the stimuli to mental representations stored in

416   LTM facilitates learning new skills. This was the basis for our hypothesis that task-specific training would yield

417   more verbalizations than traditional TA during later use of TA. Despite the intuitive logic of this hypothesis, our

418   data indicated no differences in verbalization frequency between these two TA training instructions, seemingly

419   suggesting that rich TA verbalizations have been captured through the exclusive use of traditional TA training

420   instructions in previous studies (e.g., Aitken & Mardegan, 2000; Nicholls & Polman, 2008; Samson et al., 2015)

421   and that studies that used combined traditional and task-specific TA training instructions (e.g., North et al.,

422   2011; Runswick et al., 2018; Whitehead et al., 2015) were comparable. This finding validates the large volume

423   of studies that relied exclusively on Ericsson and Simon's (1993).

424   The second aim of this study was to determine whether TA training type would impact participants'

425   subjective perceptions of TA training effectiveness. Overall, our hypothesis that task-specific TA training would

426   result in more favourable participant perceptions of training effectiveness compared to traditional TA training

427   was rejected. Analysis of the TA training checklist data revealed no significant differences between the training

428   groups. Similarly, analysis of the self-efficacy data indicated no significant differences between the training

429   groups, with both groups reporting very high levels (>86%) of self-efficacy to perform TA. Furthermore,

430   analysis of the quantitative social validation data generally revealed no significant differences in perceptions of

431   TA training, with both groups reporting that the TA training was enjoyable and effective (see Table 3). From a

432   theoretical standpoint, these data are surprising, since we expected task-specific TA training to form stronger

433   participant connections with mental golf representations of TA stored in their LTM, leading them, in turn, to

434   grasp TA more effectively and become more efficacious in using TA. These contrary findings suggest that

435   participant perceptions of TA training in studies that relied exclusively on traditional TA training instructions

436   (e.g., Aitken & Mardegan, 2000; Nicholls & Polman, 2008; Samson et al., 2015) were similar to studies using

437   combined traditional and task-specific TA training instructions (e.g., North et al., 2011; Runswick et al., 2018;

438   Whitehead et al., 2015).

439   We did find a significant training group difference in participant perceptions of instruction clarity,

440   providing some support for our contention that participants who received task-specific instructions would form

441   stronger representations of TA in golf environments in their LTM, leading them to perceive task-specific

442   instructions as having greater clarity with regard to their expectations of thinking aloud whilst playing golf. This

443   specific analysis of the qualitative social validation data provides support for our contention that task-specific

444   TA training may offer advantages over and above the traditional TA training procedures. When asked to further

445   articulate their thoughts and feelings about their training, participants offered a number of meaningful insights

446   about their experiences of learning and using TA. First, regarding the Confidence theme, participants receiving

447   traditional TA instructions reported a lack of confidence in disclosing all their thoughts as they "weren't fully

448   familiar with the task." Secondly, regarding the Task Understanding theme participants who received traditional

449   TA training said that they needed further clarification on how to do TA and how the technique can be applied to

450   golf and the task at hand. Finally, regarding the Future Support theme, participants in the traditional TA group

451   expressed the need for the training exercises to have clearer links to golf. Again, this may link to the need for

452   familiarization within the context of a given task; learning TA through task-specific instructions may be easier

453   for participants within the specific context of golf, in this instance. While studies exclusively using traditional

454 TA training instructions (e.g., Aitken & Mardegan, 2000; Nicholls & Polman, 2008; Samson et al., 2015) have

455 captured valuable verbalization data, the instruction clarity and qualitative social validation data gleaned in the

456 present study suggest that the richness of verbalizations and participant confidence may have been enhanced by

457 including task-specific training instructions. In this area, the instruction clarity and qualitative social validation

458 data therefore serve to support previous studies which have used a combination of traditional and task-specific

459 TA training approaches (e.g., North et al., 2011; Runswick et al., 2018; Whitehead et al., 2015).

460   Although this study successfully investigated TA training methods, study limitations include a lack of

461 female representation within our participant sample. Indeed, close inspection of the literature reveals a general

462 weakness in this regard in that very few studies (e.g., Arsal et al. 2016; Calmeiro & Tenenbaum, 2011; Kaiseler,

463 Polman, & Nicholls, 2010; Whitehead et al. 2018) have included female participants. There is clearly a need for

464 future research to examine TA protocols with representative female samples to better understand how TA can be

465 best trained and utilized with persons of both sexes. Furthermore, future research should consider applying the

466 methodology adopted within this current study to different domains where TA has been adopted, such as

467 medical and educational settings. A further consideration within this study could be the expertise level of the

468 participant. Although this study ensured that both groups held a very similar skill level (i.e., golf handicap), this

469 could be something for future researchers to consider when training participants to use TA.

470   Overall findings from the present study indicated no differences in verbalization or content frequency

471 and perceptions of training effectiveness between the traditional TA training protocols outlined by Ericsson and

472 Simon (1993) and the task-specific TA training protocols designed for this study. This finding lends support to

473 existing methods of TA training on which most past literature is based, in sport and exercise psychology and

474 beyond. Additionally, this study's findings provide confidence to researchers and practitioners seeking to train

475 TA effectively. However, our findings also suggest that traditional TA training protocols may be enhanced, at

476 least in terms of participants' perceptions of their clarity by the use of task-specific training exercises. In an

477 article outlining the utility of TA, Eccles and Arsal (2017) advocated the use of warm-up exercises to ensure that

478 participants are familiar with verbalizing their thoughts out loud. Indeed, Eccles and Arsal (2017) outlined

479 common pitfalls in applying the TA method, and specifically named among these: (a) allowing and encouraging

480 descriptions and explanations of thoughts, (b) failing to use warm-up exercises, (c) thinking aloud for too long,

481 and (d) possible concerns regarding participant reactivity. Given the findings from our controlled manipulations

482 of TA training methods, future TA researchers and practitioners are strongly encouraged to harness Ericsson and

483 Simon's (1993) guidelines to train TA but also to integrate task-specific training exercises to enhance participant

484   perceptions of the training process. While previous research (e.g., North et al., 2011; Runswick et al., 2018;

485   Whitehead et al., 2015) has used a combination of traditional and task-specific instructions to train TA, this is

486   the first study to provide an empirical test of advantages to this approach. While using task-specific instructions

487   may not be essential, in the context of unfamiliar tasks, they may help participants with perceived TA

488   instructional clarity.

489   Upon analysing the qualitative data gleaned from this study, it is clear that participants valued the use

490   of feedback and reiteration of principles when learning how to effectively TA. To the author's knowledge, this

491   study is the first to harness social validation methods to examine participant perceptions of TA and more

492   specifically, how to best train TA. Although it was not possible to provide spoken feedback to participants in

493   this present study without compromising experimental control, researchers and practitioners are encouraged to

494   monitor TA training progress (e.g., by using social validation methods such as TA checklists, measures of TA

495   efficacy and open questioning) to ensure all participants learn how to effectively TA before data collection

496   commences. It is important to note that previous research has used methods to monitor the learning of TA within

497   training protocols (e.g., North et al., 2011), yet similar to the TA training instructions presented in the literature,

498   the use of such learning monitoring methods has been inconsistent. Implementing more thorough TA training

499   procedures will not only enhance the participant's confidence of thinking aloud, but will also enhance the rigour

500   underpinning verbalizations, and in turn, the authenticity of verbalizations captured.

501   **References**

502   Adams, D., Cropley, B., & Mullen, R. (2016). Developing specialised youth soccer coaching qualifications: An

503   exploratory study of course content and delivery mechanisms. *International Sport Coaching*

504   *Journal*, *3*(1), 31-45.

505   Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

506   Arsal, G. (2013). *Investigating skilled and less-skilled golfers' psychological preparation strategies: The use of*

507   *a think-aloud cognitive process-tracing measure* (Doctoral dissertation, The Florida State University).

508   Arsal, G., Eccles, D. W., & Ericsson, K. A. (2016). Cognitive mediation of putting: Use of a think-aloud

509   measure and implications for studies of golf-putting in the laboratory. *Psychology of Sport and*

510   *Exercise, 27*, 18-27.

511   Balagué, N., Hristovski, R., Garcia, S., Aragonés, D., Razon, S., & Tenenbaum, G. (2015). Intentional thought

512   dynamics during exercise performed until volitional exhaustion. *Journal of Sports Sciences, 33*, 48–57.

513   Bandura, A. (1997). *Self-efficacy: The exercise of control*. Macmillan.

514   Banning, M. (2008). A review of clinical decision making: Models and current research. *Journal of Clinical*

515   *Nursing, 17,* 187-195.

516   Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under

517   pressure in sports and mental tests. *European Journal of Social Psychology*, *16*(4), 361-383.

518   Bloom, B. S., & Broder, L. J. (1950). Problem-solving processes of college students. *Supplementary*

519   *Educational Monographs*, B, 109.

520   Busse, R.T., Kratochwill, T. R., & Elliott, S. N. (1995). Meta-analysis for single-case consultation

521   outcomes: Applications to research and practice. *Journal of School Psychology, 33*, 269–

522   285.

523   Calmeiro, L., & Tenenbaum, G. (2011). Concurrent verbal protocol analysis in sport: Illustration of thought

524   processes during a golf-putting task. *Journal of Clinical Sport Psychology*, *5*(3), 223-236.

525   Cohen, J. (1994). The earth is round (p < .05). *American Psychologist, 49*, 997-1003.

526   Connors, M.H., Burns, B.D., & Campitelli, G. (2011). Expertise in Complex Decision Making: The Role of

527   Search in Chess 70 Years After de Groot. *Cognitive Science, 35*, 1567-1579.

528   Cumming, J., Hall, C., & Shambrook, C. (2004). The Influence of an Imagery Workshop on Athletes' Use of

529   Imagery. *Athletic Insight, 6*(1), 52-73.

530   De Groot, A. D. (1978). *Thought and choice in chess (Revised translation of De Groot, 1946)*. The Hague:

531   Mouton Publishers.

532   Eccles, D.W., (2012). Verbal reports of cognitive processes. In: G. Tenenbaum, R.C. Eklund, and A. Kamata,

533   eds. *Measurement in sport and exercise psychology*. Champaign, IL: Human Kinetics, 103–117.

534   Eccles, D. W., & Arsal, G. (2017). The think aloud method: what is it and how do I use it? *Qualitative Research*

535   *in Sport, Exercise and Health*, *9*(4), 514-531.

536   Elliott, D., Hansen, S., Grierson, L. E. M., Lyons, J., Bennett, S. J., & Hayes, S. J. (2010). Goal-directed aiming:

537   Two components but multiple processes. *Psychological Bulletin, 136*, 1023–1044.

538   Ellis, A. K. (2013). *Teaching, Learning, & Assessment together: Reflecting assessment for elementary*

539   *classrooms.* Routledge; New York, London.

540   Ericsson, K. A., & Simon, H. A. (1993). *Verbal reports as data*. Cambridge: MIT Press.

541   Folkman, S., & Moskowitz, J. T. (2004). Coping: Pitfalls and promise. *Annual Review of Psychology, 55*, 745-

542   774.

543   Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health*

544   *Research*, *15*(9), 1277-1288.

545   Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

546   Kaiseler, M., Polman, R. C., & Nicholls, A. R. (2013). Gender differences in stress, appraisal, and coping during

547   golf putting. *International Journal of Sport and Exercise Psychology, 11*(3), 258-272.

548   Kazdin, A. (1982). Single-case experimental designs. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of*

549   *research methods in clinical psychology* (pp. 461–490). New York: Wiley.

550   Lord, R. & Maher, K. (1991). *Leadership and information processing: Linking perception and performance*.

551   Boston, MA: Unwin Hyman.

552   O'Brien, M., Mellalieu, S., & Hanton, S. (2009). Goal-setting effects in elite and nonelite boxers. *Journal of*

553   *Applied Sport Psychology*, *21*(3), 293-306.

554   MacPhail, C., Khoza, N., Abler, L., & Ranganathan, M. (2016). Process guidelines for establishing Intercoder

555   Reliability in qualitative studies. *Qualitative Research, 16,* 198–212.

556   Mellalieu, S. D., Hanton, S., & Thomas, O. (2009). The effects of a motivational general-arousal imagery

557   intervention upon pre-performance symptoms in male rugby union players. *Psychology of Sport &*

558   *Exercise, 10*, 175-185.

559 McRobert, A.P., Causer, J., Vassiliadis, J., Watterson, L., Kwan, J., & Williams, M. A. (2013). Contextual

560   information influences diagnosis accuracy and decision making in simulated emergency medicine

561   emergencies. *BMJ Quality & Safety, 22*, 478–484.

562 McRobert, A. P., Williams, A. M., Ward, P., & Eccles, D. W. (2009). Tracing the process of expertise in a

563   simulated anticipation task. *Ergonomics, 52*, 474-483.

564 Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

565 Newton, P., & Burgess, D. (2008). Exploring types of educational action research: Implications for research

566   validity. *International Journal of Qualitative Methods*, *7*(4), 18-30.

567 Nicholls, A. R., & Polman, R. C. J. (2008). Think aloud: Acute stress and coping strategies during golf

568   performances. *Anxiety, Stress and Coping, 21*, 283-94.

569 North, J. S., Ward, P., Ericsson, A., & Williams, M.A. (2011). Mechanisms underlying skilled anticipation and

570   recognition in a dynamic and temporally constrained domain. *Memory, 19,* 155–168.

571 Page, J., & Thelwell, R. (2013). The value of social validation in single-case methods in sport and exercise

572   psychology. *Journal of Applied Sport Psychology*, *25*(1), 61-71.

573 Ptacek, J. T., Smith, R. E., Espe, K., & Raffety, B. (1994). Limited correspondence between daily coping reports

574   and restrospective coping recall. *Psychological Assessment*, *6*(1), 41.

575 Runswick, O. R., Roca, A., Williams, A. M., Bezodis, N. E., & North, J. S. (2018). The effects of anxiety and

576   situation-specific context on perceptual–motor skill: A multi-level investigation. *Psychological*

577   *Research, 82*(4), 708-719.

578 Samson, A., Simpson, D., Kamphoff, C., & Langlier, A. (2015). Think aloud: An examination of distance

579   runners' thought processes. *International Journal of Sport and Exercise Psychology*, *15*(2), 176-189.

580 Scanlan, T. K., Stein, G. L., & Ravizza, K. (1989). An in-depth study of former elite figure

581   skaters: II. Sources of enjoyment. *Journal of Sport & Exercise Psychology, 11*(1), 65-83.

582 Schmidt, R. A., & Wrisberg, C. A. (2000). *Motor learning and performance: A problem-based learning*

583   *approach*. Champaign. IL: Human Kinetics.

584 Smith, R. E., Leffingwell, T. R., & Ptacek, J. T. (1999). Can people remember how they coped? Factors

585   associated with discordance between same-day and retrospective reports. *Journal of Personality and*

586   *Social Psychology*, *76*(6), 1050.

587 Smith, B., & McGannon, K. R. (2017). Developing rigor in qualitative research: Problems

588   and opportunities within sport and exercise psychology. *International Review of Sport*

589   *and Exercise Psychology*, 101-121.

590   Stone, A., Schwartz, J., Neale, J., Shiffman, S., Marco, C., Hickcox, M., Paty, J., Porter, L., & Cruise, L. (1998).

591   A comparison of coping assessed by ecological momentary analysis and retrospective recall. *Journal of*

592   *Personality and Social Psychology, 74*, 1670–1680.

593   Taylor, S., Allsop, M. J., Bennet, M.I., & Bewick, B. M. (2017). Usability and acceptability of an electronic

594   pain monitoring system for advanced cancer: a think aloud study. *BMJ Supportive & Palliative Care,*

595   *6.*

596   Thelwell, R. C., Greenlees, I. A., & Weston, N. J. (2006). Using psychological skills training to develop soccer

597   performance. *Journal of Applied Sport Psychology*, *18*(3), 254-270.

598   Van Someren, M.W., Barnard, Y.F., & Sandberg, J.A.C. (1994). *Think Aloud method. A practical guide to*

599   *modelling cognitive processes.* London: Academic Press.

600   Ward, P., Williams, A.M., & Ericsson, K.A. (2003) Underlying mechanisms of perceptual-cognitive expertise in

601   soccer. *Journal of Sport and Exercise Psychology, 25*, S136.

602   Welford, A.T (1967). Single-channel operation in the brain. *Acta Psychologica, 27*, 5-22.

603   Welsh, J. C., Dewhurst, S. A., & Perry, J. L. (2018). Thinking Aloud: An exploration of cognitions in

604   professional snooker. *Psychology of Sport and Exercise*, *36*, 197-208.

605   Whitehead, A. E., Taylor, J. A., & Polman, R. C. (2015). Examination of the suitability of collecting in event

606   cognitive processes using Think Aloud protocol in golf. *Frontiers in Psychology*, *6*, 1083.

607   Whitehead, A. E., Taylor, J. A., & Polman, R. C. (2016). Evidence for skill level differences in the thought

608   processes of golfers during high and low pressure situations. *Frontiers in Psychology*, *6*, 1974.

609   Whitehead, A. E., Jones, H. S., Williams, E. L., Dowling, C., Morley, D., Taylor, J. A., & Polman, R. C. (2017).

610   Changes in cognition over a 16.1 km cycling time trial using Think Aloud protocol: Preliminary

611   evidence. *International Journal of Sport and Exercise Psychology*, 1-9.

612   Whitehead, A. E., Jones, H. S., Williams, E. L., Rowley, C., Quayle, L., Marchant, D., & Polman, R. C. (2018).

613   Investigating the relationship between cognitions, pacing strategies and performance in 16.1 km

614   cycling time trials using a think aloud protocol. *Psychology of Sport and Exercise*, *34*, 95-109.

615   Williams, M.A., Ford, P.R., Eccles, D.W., & Ward, P. (2011). Perceptual-Cognitive Expertise in Sport and its

616   Acquisition: Implications for Applied Cognitive Psychology. *Applied Cognitive Psychology, 25*, 432–

617   442.

618   Wolf, M. W. (1978). Social validity: The case for subjective measurement or how applied behavior

619   analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203–214.

620

Table 1. Content summary of the TA training videos.

Content	Traditional TA training	Task-specific TA training
Introduction	TA background information provided (Ericsson & Simon, 1993)	
TA level	Instructions on how to TA – level 2 (Ericsson & Simon, 1993)	
Authenticity	Instructions were provided to emphasise process of TA	
TA training	Exercises based on Ericsson and Simon (1993):	Three scenarios were used to stimulate TA. Participants were asked to TA their thoughts on a hypothetical par 5 golf hole for their:
	4 x alphabetical problems solving tasks	
	5 x counting the number of dots on a page	Tee shot
		Fairway (second) shot
	2 x general problem solving tasks	
		Greenside (third) approach shot

| Recap | Participants were asked to recall the key principles of TA |
| | Researcher reminded participant of principles missed |

TA practice | 3 x trials on the golf task whilst thinking aloud

Note: Training videos are available on request.

621

622

Table 2. TA coding framework.

Theme	Description
Gathering information	Refers to participants' search for relevant characteristics of the environment (e.g., "there's a break left," "it is mostly uphill").
Planning	Refers to the definition of actions or strategies to reach a goal (e.g., "aim two cups right," "hit firm at the hole").
Mental readiness	Refers to psychological preparation for the task (e.g., "you know you can do this," "concentrate on this").
Technical instruction	Refers to specified technical aspects of the motor performance (e.g., "arms bent," "feet are parallel").
Description of outcome	Refers to what had happened in terms of process or evaluation of the action (e.g., "[the ball] flew that by," "it broke at the end," "good putt").
Diagnosis of outcome	Refers to the reasons for the observed outcome (e.g., "I didn't hit hard enough," "too firm").
Reactive comments	Refers to verbalisations referring to reactive comments to performance (e.g., "This hole is not working for me!" "Oh, goodness . . . it should have gone in!").

Adapted from Calmeiro and Tenenbaum (2011) and Whitehead et al. (2016).

623

624

625

626

627

628

Table 3. Means and standard deviations of themes verbalised, task commitment scores, TA training checklist scores, self-efficacy scores and social validation scores between the tradition TA training group and the task-specific TA training group.

	Traditional TA training		Task-specific TA training	
	M	*SD*	*M*	*SD*
Themes verbalised				
Gathering information	17.10	7.78	15.80	7.41
Planning	24.10	7.53	20.50	7.47
Mental Readiness	2.70	5.90	5.60	6.27
Reactive Comments	1.30	3.13	3.00	4.24
Description of Outcome	17.70	4.47	15.50	5.40
Diagnosis of Outcome	5.20	4.47	5.60	3.16
Technical Information	3.80	6.62	6.00	5.01
Task commitment				
Post practice	85.50	12.12	94.00	9.66
Post TA training	97.00	4.83	95.00	9.72
TA training checklist	2.80	1.69	3.20	1.23
Self-efficacy				
Post practice	86.00	21.19	89.50	9.56
Using TA in the future	86.00	13.50	89.00	9.94

Social validation

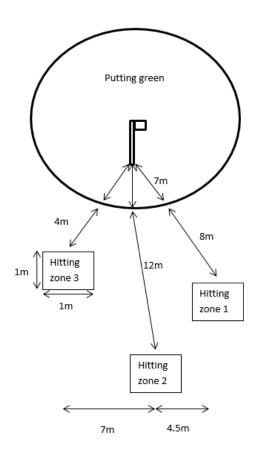|				
Enjoyment of using TA	5.80	1.81	6.60	0.70
Clarity of instructions	6.10	1.20	7.00	0
In-video TA training task effectiveness	4.90	2.13	6.30	1.06
Physical TA practice trials effectiveness	6.70	0.95	6.30	1.25
Overall TA training effectiveness	6.20	1.23	6.40	0.97

Note: M = Mean, SD = Standard Deviation.

630

631   **Figure 1.** Schematic representation of the golf task.

632

633