Available online at www.scholarsresearchlibrary.com

Scholars Research Library

European Journal of Sports & Exercise Science, 2019, 7 (1): 166-171

(http://www.scholarsresearchlibrary.com)

ISSN:2278-005X

Examining the Suitability of Traditional and Task-Specific Training for Think Aloud Protocol

Amy Whitehead*

Health and Community, Liverpool John Moores University, UK

***Corresponding Author:** Amy Whitehead, Faculty of Education, Health and Community, Liverpool John Moores University, UK, Tel: 01512314613; E-mail: A.E.Whitehead@ljmu.ac.uk*

ABSTRACT

Introduction: Think Aloud protocol (TA) is a method used to capture conscious cognition in a wide range of domains. However, the methods used to train TA protocol with participants appear to be inconsistent in that a mixture of both traditional guidelines and task-specific examples have been used. Therefore, the aim of this study was to examine how best to train the TA protocol.
Method: 20 competitive golfers were recruited as participants. 10 participants were randomly assigned to a traditional TA training group which comprised of TA training guidelines forwarded by Ericsson and Simon. 10 participants were randomly assigned to the task-specific training group where participants were familiarised with TA using task-specific examples. Following training, all participants performed a golf task and were asked to TA. TA audio was transcribed verbatim and analyzed using a deductive framework. A variety of social validation measures were collected to assess participant perceptions of TA training. Results: Overall findings revealed no significant differences in the frequency and type of verbalizations between the traditional and task-specific TA training groups. However, social validation findings indicated that participants in the task-specific training group reported more favorable perceptions and found the training significantly clearer than the traditional training group.
Conclusion: This study provides support for researchers using TA and encourages the use of Ericsson and Simon's training guidelines, however with the addition of task-specific examples to increase familiarity of using TA, which in turn, will facilitate more reliable and accurate cognition data.

Keywords: Think aloud, Cognition, TA training, Exercise.

INTRODUCTION

Examining the suitability of traditional and task-specific training for Think Aloud protocol. Think Aloud (TA) protocol or verbal reports [1] have been used for decades to capture in event problem solving and decision-making data [2]. TA has been used in a variety of domains such as medical settings relating to pain management [3], surgery [4], nursing [5], teaching [6] and within the sport to capture in-performance cognitions [7,8]. The terminology surrounding TA and verbal reports has been used interchangeably, with some researchers using the term TA [9] and others using terms relating to the specific time of the reports including immediate retrospective verbal reports [4]. For this paper, we will use the term TA as an umbrella term when discussing the literature and more detail on how TA is used will be provided throughout the method section. TA involves a performer verbalizing his or her thought process continues as they are performing. Ericsson and Simon [1] proposed three levels of TA. Level 1 verbalization is simply the vocalization of inner speech where the individual does not need to make any effort to communicate his or her thoughts. Level 2 verbalization involves the verbal encoding and vocalization of an internal representation that is not originally in verbal code (e.g., verbal encoding and vocalization of scents, visual stimuli, or movement). With this level of verbalization, only the information that is in the participant's focus is to be verbalized. Level 3 verbalization requires the individual to explain his or her thoughts, ideas, hypotheses, or motives [1]. For example, explaining why a certain medical procedure should be conducted or explaining why a certain shot or club is selected in golf. It is important to note that the majority of the literature using TA has opted for Level 2 due to the potential limitations of Level 3 in that it may interfere with the natural thought process [1]. The benefit of using TA is that it allows data to be captured within real time and therefore reduces the risk of memory decay and retrospective bias [10-13]. Furthermore, TA allows researchers to identify potential differences in perceptual-cognitive processes between expertise levels [14]. For example, evidence of this domain-specific knowledge and mental representation can be demonstrated by the early work of De Groot [15]. This work demonstrated that when asking master and intermediate chess players to reconstruct the locations of chess pieces after viewing the board for only a few seconds, master chess players were much superior to intermediates. However, when the pieces were presented randomly, the two groups did not differ. What is evident here is that, over time and practice the higher level chess players have been able to store thousands of chunks of chess related information (a chunk was defined as a sequence of pieces with between piece intervals of less than two seconds) in their Long Term Memory (LTM) which they are able to retrieve and as a result are more familiar with and remember more positions after only a few seconds of viewing the board. Research using TA has identified cognitive differences between various levels of performers in a wide range of domains. Within medical research, McRobert et al. [4] found that skilled physicians demonstrated higher diagnostic accuracy and selected better quality options during diagnostic reasoning compared to less skilled physicians. In chess, the research identified how Grandmaster players search more quickly and have superior pattern recognition than players of lower level [16]. Furthermore, Whitehead, Taylor, and Polman [17] found that lower skilled golfers thought processes focused more on the technical mechanics of their performance, whereas higher skilled golfers thought processes focused more on planning skill execution. The information gleaned from these studies may enable practitioners to identify potential flaws in cognitive strategies and provide key information for the implementation of future interventions to enhance performance.

Despite the importance placed on TA training by Ericsson and Simon [1], these researchers provided general instructions which included one simple mental arithmetic task and one problem-solving task designed to familiarise the participants with the TA technique. Specifically, Ericsson and Simon [1] stated: “Good, before we turn to the real experiment, we will start with a couple of practice problems. I want you to talk aloud while you do these problems. First, I will ask you to multiply two numbers in your head. So talk aloud while you multiply 24 times 34. Good! Now I would like you to solve an anagram. I will show you a card with scrambled letters. It is your task to find an English word that consists of all the presented letters. For example, if the scrambled letters are KORO, you may see that these letters spell the word ROOK. Any questions? Please “talk aloud” while you solve the following anagram! <NPEPHA=HAPPEN>.”

Adaptations of these warm-up tasks [18] have been used in previous research that has utilized the TA procedure [19,20]. However, these tasks are non-task specific and it is unknown to what extent participants feel that these tasks fully equip them with the ability and confidence to effectively perform TA. Indeed, Van Someren, Barnard and Sanberg [21] highlight the importance of aligning the training task to the target task, or as they state “in general it is wise to look for a task which is not too different from the target task”.

When learning a new skill, domain specificity is extremely important, especially from an information processing perspective. When a new skill is being processed, the body (one or more of the sense organs) identifies the task/ stimuli and a response is selected, prepared and initiated. This process involves internal memorial representations [15,22]. During this new activity/engagement with the new stimuli, it is coded and sent to the brain, where it is identified as new or familiar according to its similarity to the representations already stored in LTM. Lord and Maher [23] provide a simplistic way of explaining how a task is performed from an information processing perspective. They propose a view that emphasizes the energy required to perform a task. More specifically, the number of tasks that can be performed concurrently is limited by the combined amount of energy that tasks consume [24,25]. The energy requirements needed to perform a task depends on how well the task has been practiced. Therefore, novel tasks require much more energy or attention (controlled processing) while well-rehearsed tasks require fewer attentional demands (automatic processing). It could be argued that if a task such as learning TA is closely linked to the domain it is being performed, then the energy to perform the task within this familiar environment may be less than if it is not domain specific. Therefore, when learning TA in a specific environment, it could be intuitive to predict that being trained to use TA with task-specific examples may make the process easier to grasp, given that it allows for connections to be made with task-specific representations already stored in LTM. In an effort to supplement the traditional TA training methods recommended by Ericsson and Simon [1], some research has used task-specific warm-up tasks to better familiarise participants with TA. North, Ward, Ericsson, and Williams [26] provided the following information; “several domain-specific examples were included as part of the training protocol. The training session included instruction and practice at thinking aloud, and retrospectively reporting these thoughts using a range of generic problems and task-specific video-based scenarios”. In Arsal, Eccles and Ericsson’s

[27] study, participants "practiced thinking aloud while putting twice over 89 cm". Runswick, Roca, Williams and Bezodis [28] stated: "training included instruction on thinking aloud and giving immediate retrospective verbal reports by solving a range of generic and domain-specific tasks". Similarly, Calmeiro and Tenenbaum's [29] second phase of TA training "consisted of verbalisation practice while putting" and Whitehead et al.'s [15] TA protocols were "adapted to golf putting based upon the guidelines set out by Ericsson and Simon [1] and Nicholls and Polman [20]." Despite Whitehead et al. [30,31], providing participants with task (cycling) specific video material prior to data collection, it is not entirely clear what this involved. Although it is positive to see some task-specific TA training being implemented, it could be argued that TA training procedures could be strengthened by more consistent use of task-specific warm-up tasks to ensure replicability in the training of TA. Enhancing the specificity of TA instructions could lead to a number of favorable outcomes. Firstly, it could increase the reader's ability to fully understand how TA is trained, and in turn, enable researchers and practitioners to replicate the administered TA training procedures. At present, it could be argued that the information reported in the literature is limited and thus may not afford accurate replication of protocols. Indeed, Samson, Simpson, Kamphoff, and Langlier [32] conceded that a limitation of their study was the "non-sport nature of the warm-up tasks" and have encouraged researchers to examine the effectiveness of TA training protocols. Secondly, it may increase the participant's ability to effectively learn how to TA, which in turn, could enhance the quality of verbalizations captured.

Given the importance placed on upholding rigor in the data capture process [1], greater understanding is needed concerning the precise procedures utilized to train TA. Consequently, further research is required to examine how best to train TA, and in turn, what impact TA training might have on athlete verbalizations.

In order to shed light on the effectiveness of TA training procedures, it would appear intuitive to ask athletes their opinions. To the author's knowledge, previous research has not explicitly examined how a participant perceives the TA training process. Traditionally, social validation procedures have been used to measure participant perceptions and satisfaction of an intervention [33,34]. Consequently, it could be argued that social validation affords a method to examine athlete perceptions of the respective components of TA training (e.g., clarity of verbal instructions, the effectiveness of training exercises), and in turn, the effectiveness of TA training methods. Further research is therefore warranted to examine methods of training TA in an effort to shed light on how best to train TA. This may afford a more consistent approach to using TA, which in turn, may lead to more in-depth understanding of cognitive processes in experts across domains. Due to the exploratory nature of this paper and that no research has explicitly examined differences in how TA is trained, this study aimed to examine the impact of traditional and task-specific TA training procedures on cognitive processes and perceptions of training effectiveness. Given that more positive perceptions of TA training (e.g., the confidence of using TA) could be associated with a higher willingness to verbalize one's cognition, it was hypothesized that the task-specific TA training may result in significantly more verbalizations when compared to the traditional TA training. Given that task-specific training may promote greater storage of contextual information in the LTM, it was hypothesised that the task-specific training would result in more favourable perceptions of TA training effectiveness when compared to the traditional TA training.

METHOD

Participants

Participants were 20 golfers from a golf club in the South of England. Participants were split into two equal skill groups consisting of a traditional TA training group (n=10; 6 males, 4 females; age: M=42.7 years, SD=11.8; handicap: M=13.1, SD=10.4) and a task-specific TA training group (n=10; 10 males, 0 females; age: M=43.0 years, SD=14.2; handicap: M=12.5, SD=10.3). Participants in the traditional TA training group had an average of 11.2 (SD=9.6) years competitive playing experience, played at least once per week and had played an average of 19.6 (SD=11.8) competitions in the 12 months leading up to their participation in the study. Participants in the task-specific TA training group had an average of 19.7 (SD=9.0) number of years competitive playing

experience, played at least once per week and had played an average of 11.1 (SD=7.1) competitions in the 12 months leading up to their participation in the study. All participants identified their ethnicity as white British. Institutional ethical approval was secured and informed consent was obtained from all participants prior to participation.

MATERIALS

TA training videos

All participants used their own clubs and balls to perform the golf task, which was conducted on a practice green at the home golf course of the participants. A Sony HXR-NX30N camcorder with radio microphone was used to record verbalizations. The mini radio microphone was attached to the participant's collar, and a wire placed inside the shirt connecting to the recording device which was placed in the participant's pocket.

The stimuli used in this experiment were two TA training videos, each consisting of visual and verbal instructions on how to perform TA (Table 1 for the content summary of each video). The purpose of the videos was to provide participants with an understanding of how TA works so they could competently perform the technique. In line with Ericsson and Simon's [1] guidelines, both videos provided identical instructions to train participants in performing TA. Example instructions included "think aloud involves you saying out loud everything that you are thinking as you are performing the task" and "it is important that you think aloud all your thoughts as best as you can during that time". Given that this study aimed to examine level 2 TA, both training videos included instructions to promote level 2 TA and deter level 3 TA. In accordance with Ericsson and Simon's [1] guidelines, the videos stated "I don't want you to try to plan out what you say or try to explain to me what you are saying". In order to promote authentic projection of thoughts [1], both videos instructed participants to "just act as if you are alone speaking to yourself". To ensure that participants were performing TA throughout the golf task [1], the videos stated "It is most important that you keep talking. If you are silent for any long period of time, I will ask you to talk." Further instructions were also written specifically for this study. These included "we are interested in knowing your thoughts as they come to mind during the golf task", "this includes the thoughts you have in the lead up to hitting the ball, whilst the ball is in motion, after the ball has come to rest, and as you walk to play your next shot" and "everything you say is confidential-the researcher will not judge your thoughts and please use swear words if you feel necessary". It is important to note that participants were instructed to refrain from verbalizing during skill execution to reduce possible interference with motor movement [19]. The remainder of the videos consisted of the participant's TA training treatment; traditional TA training or task-specific TA training. The training exercises used in the traditional TA training video were based on the recommendations of Ericsson and Simon [1] and consisted of three different groups of tasks: a) four alphabetical problems solving tasks (e.g., what is the fourth letter after H); b) five counting the number of dots on a page; c) two general problem-solving tasks (e.g., name two vegetables that begin with the letter C). These training exercises have been used in a number of previous research studies [8,17,20,32].

The exercises used in the task-specific training video were developed for this study and consisted of three different golf scenarios: 1) Tee shot on a par 5; 2) fairway (second) shot on a par 5; 3) greenside approach (third) shot over a bunker. For the first scenario, the following information was provided: "It 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 the ball very well and scoring very well in the lead up to this competition. It is a reasonably warm summer's day and the course is firm and playing fast. The weather is overcast and there is a strong wind against." For the second scenario, the following information was provided: "You are now playing your second shot on the same hole in the monthly medal. Again, you have been striking the ball very well and scoring very well in the lead up to this competition. It is a reasonably warm summer's day and the course is firm and playing fast. The weather is overcast and there is a strong wind against. The pin is cut back right. Your ball is in the middle of the fairway and lying very nicely. Your ball is marked by the white ball on the right diagram". For the third scenario, the following information was provided: "You are now playing your third shot on the same hole in the monthly medal. Your short game has been poor in the lead up to this competition. It is a reasonably warm summer's day and the greens are playing firm and fast. The weather is overcast and there is a strong wind against. The flag is cut back right. Your ball is lying poorly in the left rough-marked by the white ball on the right diagram."

At the end of each description for the respective scenarios, participants were instructed: "please use the information in the diagrams and tell us your thoughts on how you would play this shot." At this moment, two diagrams appeared on the video to help facilitate TA. The diagram on the left provided a bird's eye view of the hole and the yardages to and from its respective features (e.g., yardage to the bunker from the tee). The diagram on the right was a first-person view of the hole (albeit from an elevated position) and represented the information a golfer would gain whilst performing on a golf course. Once the participants received their TA training treatment, both groups were instructed to complete the TA training checklist to assess how well the participant had learned the requirements of TA. Finally, all participants were instructed to have three practice trials on the golf task while verbalizing to familiarise themselves thinking aloud. The traditional TA training video was 4:47 minutes in duration and the task-specific TA training video was 4:14 minutes in duration.

Table 1: Content summary of the TA training videos.

Content	Traditional TA training	Task-specific TA training
Introduction	TA background information provided [1]	
TA level	Instructions on how to TA-level 2 [1]	
Authenticity	Instructions were provided to emphasize the process of TA	
TA training	Exercises based on Ericsson and Simon (1993): × alphabetical problems solving tasks × counting the number of dots on a page 2 × general problem-solving tasks	Three scenarios were used to stimulate TA. Participants were asked to TA their thoughts on a hypothetical par 5 golf hole for their: Tee shot Fairway (second) shot Greenside (third) approach shot
Recap	Participants were asked to recall the key principles of TA Researcher reminded participant of principles missed	
TA practice	3 × trials on the golf task whilst thinking aloud	
Note: Training videos are available on request		

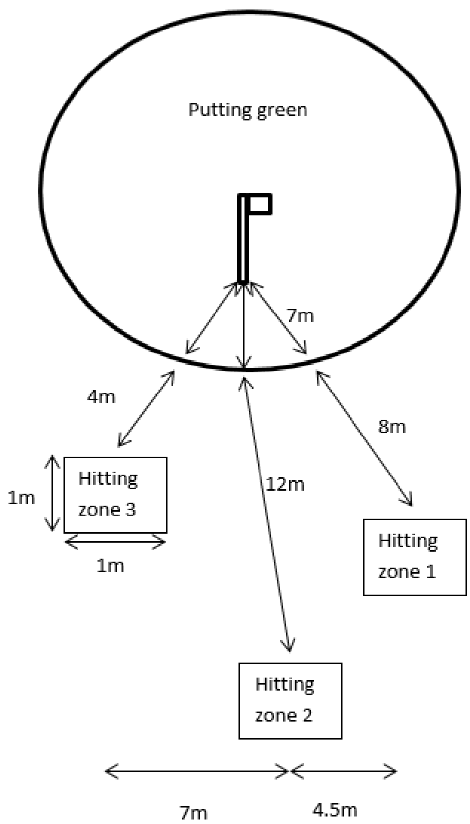

Figure 1: Schematic representation of the golf task.

The golf task: The golf task was specifically designed for this study as a means to facilitate authentic short game golf shots (i.e., chipping and putting) that golfers would typically face during a round of golf. Given that every shot is different whilst playing a round of golf, three different hitting zones were used (Figure 1). Hitting zone 1 was positioned on an up-hill lie 15 meters from the hole and exhibited an incline. Hitting zone 2 was positioned on flat lie 19 meters from the hole and exhibited an incline. Hitting zone 3 was positioned on a side-hill lie (ball below participant's feet for a right-handed golfer) 11 meters from the hole and exhibited a decline. All hitting zones were located on shortly mown grass and participants were permitted to place their ball within one meter squared area. The speed of the green was measured on a Stimpmeter. The total amount of feet the ball rolls from the Stimpmeter gives an approximation as to the pace of the green. The green measured at an average of 9 on the Stimpmeter. Participants were required to hit the ball in the hole in as few shots as

possible and were allowed to select which club to use. In order to enhance the ecological validity of the task, a series of pressure manipulations were enforced [35]. Participants were informed that they would be entered into a competition where the lowest score would receive three premium golf balls. Participants were also informed that their performance scores (i.e., amount of shots taken) would be published on a leaderboard which would be readily available for other participants to see before they performed their trials. Indeed, participants were informed of other participants' scores before performing to facilitate the comparative and evaluative nature of the task.

Measures

TA protocol: Level 2 verbalizations were recorded during the golf task. Participant verbalizations were transcribed verbatim.

Task commitment: Task commitment was measured to determine the level of engagement with the task and to determine if there were differences in task engagement between TA training groups. In accordance with research by Arsal [36], the following item was used: "How committed were you to the task while performing?" Participants were instructed to rate their commitment on a scale with 10% increments ranging from 0% (not at all) to 100% (very much).

TA training checklist: The TA training checklist was designed specifically for this study and required participants to recall seven key components to successfully perform TA: 1) confidentiality of verbalizations; 2) all thoughts are to be spoken; 3) refrain from explaining your thoughts; 4) use TA before and after your shot; 5) refrain from verbalizing during skill execution; 6) periods of silence will result in being prompted; and 7) swearing is permitted. This was used to assess how well the participants had learned the requirements of TA.

Self-efficacy: Self-efficacy for thinking aloud was measured to determine participants' belief in thinking aloud whilst performing the golf task. In accordance with Bandura's [37] recommendation, participants indicated the strength of their self-efficacy for thinking aloud concurrently to performing the golf task by responding on a one-item Likert-type scale with 10% increments ranging from 0% (not at all confident) to 100% (completely confident).

TA social validation: Social validation procedures are suggested to strengthen the external validity of technical and practical action research by offering a personal insight into the intervention through the experiences of the participants [38]. In accordance with Page and Thelwell's [39] guidelines, quantitative social validation questions were used in an effort to better understand participants' experiences of using TA. Participants were asked the following questions: 1) Did you enjoy the think aloud training? (with responses ranging from 1=not at all enjoyable to 7=very enjoyable); 2) How clear were the instructions in the think-aloud training video? (with responses ranging from 1=unclear to 7=very clear); 3) With regards to helping you learn to think aloud, how effective were the think aloud practice tasks in the training video? (with responses ranging from 1=not at all effective to 7=very effective); 4) With regards to helping you learn to think aloud, how effective were the physical TA practice trials? (with responses ranging from 1=not at all effective to 7=very effective); 5) Overall, how effective did you think the training was in preparing you to think aloud during the golf task? (with responses ranging from 1=not at all effective to 7=very effective). Participants were also asked the following qualitatively orientated open questions: 1) Is there anything that you would add to the think-aloud training? 2) Do you have any further comments regarding the TA training?

PROCEDURE

Pilot study

The pilot study consisted of two moderately skilled golfers. The golfers had handicaps of 7 and 10 and accumulated 12 and 10 years of competitive playing experience, respectively. Both golfers completed the entirety of the experimental procedure. One golfer received the traditional TA training and one golfer received the task-specific TA training. Based on the feedback, participants were confident they could verbalize whilst performing the golf task and that the equipment did not hinder their performance. Participants stated that the golf task provided a realistic task which translated well to the golf course.

Experimental procedure

Prior to conducting the experimental procedure, participants Completed a demographic questionnaire and informed consent form. All participants performed a total of 15 practice trials comprising of five trials from the three different hitting zones to familiarize themselves with the demands of the golf task (Figure 1). Trials were performed in sequential order (hitting zone 1, hitting zone 2, hitting zone and so forth) to decrease the likelihood of boredom. It was decided from the pilot testing that 15 practice trials were appropriate as this provided sufficient time to warm-up and familiarise oneself with the practice green without being too laborious. Each trial on the golf task required the participant to place the ball in the hitting zone, perform their usual pre-performance routine, hit the approach shot as if they would do on the golf course, walk up to where the ball finished, perform their usual pre-performance putting routine and attempt to putt the ball into the hole in as few shots as possible. Participants were permitted to change their clubs accordingly. Participants then received their TA training video using an Apple iPad and Sony MDR ZX660AP headphones. Participants were required to complete the TA training checklist. In order to give participants the opportunity to practice using TA whilst performing, three practice trials were given. During this time, the researcher ensured the participant was competently using TA in line with the instructions given in the training video. Participants were then asked to rate their level of self-efficacy in thinking aloud whilst performing the golf task. Participants completed a series of nine trials on the golf task whilst thinking aloud. Participants were reminded to TA throughout the nine trials and were told that if they were silent for a period longer than five seconds, they would be asked to resume thinking aloud. Although previous research has used 20 second [20] and 10 second (e.g., Whitehead et al., [17] prompt durations to ensure the occurrence of verbalizations, the pilot study revealed the need to use a shorter duration due to the relatively short gaps between skill execution on the golf task. The researcher walked to the side of the participants (approximately five meters) during the golf task and there was no communication beyond the investigator reminding the participants to continue thinking aloud and what zone to hit from next [20]. Other than the presence of the researcher, each participant performed alone. Participants were asked to rate their level of commitment with the golf task after the 15 practice trials and after the nine trials of thinking aloud. At the completion of the think-aloud trials, participants completed the social validation questions and the self-efficacy scale (to assess the efficacy of using TA in the future).

DATA ANALYSIS

Quantitative data gleaned from task commitment scale, TA training checklist, self-efficacy scale, and the social validation questions were analyzed using SPSS Statistics 23. Given that the data were normally distributed, a series of independent samples t-tests were conducted to examine differences between the traditional TA training group and the task-specific training group TA verbalizations were transcribed verbatim and were subjected to the line by line content analysis. Given the non-anticipatory nature of the golf task used in this present study, a golf-specific adapted framework from Calmeiro and Tenenbaum [29] and Whitehead et al. [8] was used to code the verbalizations (Table 2). The first author analyzed a 10% sample of the data and an interrater reliability agreement of 85% was found [40]. Discussions regarding the discrepancies between the remaining 15% took place and both authors came to an agreement. Given that the data were non-normally distributed, Mann-Whitney U tests were used to analyze the differences in themes verbalized between the traditional TA training group and the task-specific training group. Cohen's [41] d effect sizes were calculated to establish the magnitude of differences between the traditionally trained and task-specific trained participants.

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 verbalizations 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 [29] and Whitehead et al. [8]	

The qualitative social validation data were independently analyzed by the second author to ensure content familiarity. Inductive content analysis was used to analyze this data [42] Following previous research which has investigated participants perceptions of using TA [31], inductive reasoning was employed with a view of allowing themes to be generated from the data. As a result of the inductive content analysis process, three themes were generated from this data. To ensure for rigor, the lead author acted as a critical friend to ensure the data collection and analyses was plausible and defendable [43].

RESULTS

Content of TA data

A comparison of the total number of verbalisations between traditional (n=720, M=71.9, SD=20.70) and task-specific (n=719, M=72.00, SD=21.03) TA training found no significant difference: U=48.50, p=0.91, d=0.00. A series of Mann-Whitney U tests were conducted to investigate the content of the verbalisations of the traditional training group and the task-specific training group (Table 3 for descriptive statistics). No significant differences were found when comparing frequency of themes verbalised between the traditional training group and the task-specific training group: gathering information: U=49.00, p=0.49, d=0.17; planning: U=34.50, p=0.24, d=0.48; mental readiness: U=32.40, p=0.17, d=0.48; reactive comments: U=33.50, p=0.17, d=0.47; description of outcome: U=37.50, p=0.34, d=0.44; diagnosis of outcome: U=41.00, p=0.50, d=0.10; technical information: U=33, p=0.20. d=0.37.

Task commitment

Independent samples t-tests showed no significant difference in both post practice commitment check scores, t(18)=1.734, p=0.100, d=0.775, and commitment check scores post TA trials, t(18)=-0.583, p=0.567, d=0.260, between the traditional TA training group and the task-specific TA training group.

TA training checklist

An independent samples t-test showed no significant difference, t(18)=0.606, p=0.552, d =0.271, in the amount of TA instructions recalled between the traditional TA training group and the task-specific TA training group.

Self-efficacy

An independent samples t-test showed no significant difference in both self-efficacy of using TA post practice trials, t(12.519)=0.476, p=0.642, d=0.213, and self-efficacy scores in using TA in the future scores, t(18)=-0.566, p=.578, d=0.253, between the traditional TA training group and the task-specific TA training group.

TA social validation–quantitative

An independent samples t-test showed a significant difference, t(9)=2.377, p=0.041, d=1.063, in perceptions of instruction clarity between the traditional TA training group (M=6.10, SD=1.20) and the task-specific TA training group (M=7.00, SD=0). No significant differences were found when comparing the remaining social validation questions between the traditional TA training group and the task-specific TA training group (Table 3 for descriptive statistics): enjoyment of using TA scores: t(11.618)=1.302, p=.218, d=0.582; effectiveness of the in-video TA training tasks: t(13.190)=1.860, p=0.085, d=0.831; effectiveness of the physical TA practice trials: t(18)=-0.805, p=0.431, d=0.360; overall TA training effectiveness: t(18)=-0.405, p=0.691, d=0.181.

Table 3: Means and standard deviations of themes verbalized, 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	
Themes verbalized	**M**	**SD**	**M**	**SD**
Gathering information	17.1	7.78	15.8	7.41
Planning	24.1	7.53	20.5	7.47
Mental readiness	2.7	5.9	5.6	6.27
Reactive comments	1.3	3.13	3	4.24
Description of outcome	17.7	4.47	15.5	5.4
Diagnosis of outcome	5.2	4.47	5.6	3.16
Technical information	3.8	6.62	6	5.01
Task commitment				
Post-practice	85.5	12.12	94	9.66
Post TA training	97	4.83	95	9.72
TA training checklist	2.8	1.69	3.2	1.23
Self-efficacy				
Post-practice	86	21.19	89.5	9.56
Using TA in the future Social validation	86	13.5	89	9.94
Enjoyment of using TA	5.8	1.81	6.6	0.7
Clarity of instructions	6.1	1.2	7	0
In-video TA training task effectiveness	4.9	2.13	6.3	1.06
Physical TA practice trials effectiveness	6.7	0.95	6.3	1.25
Overall TA training effectiveness	6.2	1.23	6.4	0.97
Note: M=Mean, SD=Standard Deviation				

Social validation-qualitative

Analysis of verbal responses revealed three main themes within this data; Confidence, Task Understanding, and Further Support. Within these themes, it was apparent that the traditional training group and the task-specific training group exhibited different thoughts about their training.

Confidence

Both the traditional training group and the task-specific training group reported being confident in the use of TA. However, within the traditional training group, some participants reported that they may not have always verbalized everything that they would be thinking as they were not comfortable with disclosing their thoughts. Participant 9 (traditional TA group) stated: "something's I did not say as I was not fully familiar with the task and not used to blurting things out". Conversely, the task-specific group exhibited confidence in their ability to follow the training and use TA. For example, participant 4 (task-specific TA training group) stated: "they were good because they played as a scenario that I could think and apply it to my own ability."

Task understanding

Both the traditional and the task-specific training groups reported a general consensus that they understood the training tasks that they were given. However, within the traditional training TA group, some participants reported losing their way and questioned some of the TA training tasks. For example, participant 11 (traditional TA training group) stated "I lost my way a bit through the training" and participant 13 stated, "the dots were effective, but the other parts of the task, not so much". Whereas participant 10 (task-specific TA training group) reported "Yeah it just gets you into the mode, er, of thinking, with a prompt here or there if I wasn't or when I should be doing so it was good. Very helpful." Moreover, participant 14 (task-specific TA training) stated "Erm,

I just thought it was pretty simple, it just wasn't too complicated as well and I was clear about what I had to do".

Future support: Both TA training groups suggested that supplementary support would ai their ability to proficiently use TA. However, the recommendations provided were different depending on the type of training received. Participants in the task-specific training group generally reported that they would like to have had more support in using TA in the future and be reassured that the process requires all thoughts to be verbalized. For example, participant 10 (task-specific TA training group) stated: "I'd probably like to do more of it as it was a really good learning tool". Furthermore, participants in the task-specific TA training group also reported that they would like to be reminded more within the initial training that all thoughts, no matter how obscure should be verbalized. For example, participant 2 (task-specific TA training group) stated: "I would have liked to be re-assured more that even strange thoughts should be spoken out loud". Conversely, participants in the traditional TA training group reported that they needed more feedback as to whether they were doing TA properly and would have liked more exercises linking TA to the golf environment. For example, participant 11 stated "I would have liked more comments around the process of TA", and participant 2 (both traditional TA training group) specified "it could have been more golf related". Furthermore, participant 15 (traditional TA training group) expressed the need for having a clearer link to golf by stating "I would have liked more feedback in terms of how the thinking aloud will then relate to golf and if I'm doing it properly".

DISCUSSION

The first aim of this study was to investigate whether TA training type would impact verbalization frequency. The hypothesis that the task-specific TA training would result in significantly more verbalizations when compared to the traditional TA training was rejected. Findings indicated no significant differences in the categories of verbalizations between the traditional TA training group and the task-specific TA training group (Table 3). According to information processing theorists [15,22], the familiarity of the stimuli to representations stored in LTM facilitates learning new skills, hence why it was hypothesized that task-specific training may yield more verbalizations during the use of TA. Therefore, despite it being intuitive to assume that being trained to use TA with task-specific examples may strengthen familiarity of the new TA stimuli in the LTM and potentially facilitate higher levels of confidence in using TA, the data appears to indicate that there are no differences in the frequency of TA verbalizations between the traditional TA training instructions and the task-specific instructions. This finding appears to suggest that the richness of TA verbalizations captured exclusively using traditional TA training instructions in previous studies (e.g., Aitken and Mardegan, [20,32] may have been equitable to those studies which used a combination of traditional and task-specific TA training instructions [17,26,28). Given the volume of studies which have exclusively used Ericsson and Simon's [1] general guidelines for training TA, this finding serves to validate the traditional TA training approach and provides researchers and practitioners with confidence in the richness of the TA verbalisations captured The second aim of this study was to investigate whether TA training type would impact perceptions of TA training effectiveness. Overall, the hypothesis that the task-specific TA training would result in more favorable perceptions of training effectiveness when compared to the traditional TA training was rejected. Analysis of the TA training checklist data revealed no significant differences between the training groups. Analysis of the self-efficacy data indicates no significant differences between the training groups with both groups reporting very high levels (>86%) of self-efficacy to perform TA. Analysis of the quantitative social validation data generally revealed no significant differences in perceptions of TA training with both groups reporting that the TA training was enjoyable and effective (Table 3). From a theoretical standpoint, these data are surprising in that one would predict the task-specific TA training group to have formed stronger connections with golf representations of TA stored in their LTM, and in turn, grasped TA more effectively and been more efficacious in using TA. These findings appear to suggest that participant perceptions of TA training in those studies which exclusively used traditional TA training instructions [20,32] may be similar to those studies which used a combination of traditional and task-specific TA training instructions [17,26,28].

However, findings showed a significant difference in perceptions of instruction clarity between the traditional TA training group and the task-specific TA training group. These data, therefore, provide some support for the contention that participants who received task-specific instructions may have formed stronger representations of TA in golf environments in their LTM and thus deemed the instructions to exhibit greater clarity regarding the expectations of thinking aloud whilst playing golf. Analysis of the qualitative social validation data provides support for the contention that the task-specific TA training may offer advantages over and above the traditional TA training procedures in that some differences between the TA training groups were observed. When asked to further articulate their thoughts and feelings about their training, participants offered a number of meaningful insights about their experiences of learning and using TA. Firstly, the Confidence theme was characterized by differences in belief as to the relevance of training exercises. For example, participants receiving traditional training reported a lack of confidence in disclosing all their thoughts as they "weren't fully familiar with the task". Secondly, the Task Understanding theme was characterized by differences in the need for clarification of the purpose of training exercises. For instance, participants who received tradition TA training said that they needed further clarification on how to do TA and how the technique can be applied to golf and the task at hand. Finally, the Future Support theme was characterized by differences in supplementary guidance for effectively using TA in the future. For example, participants in the traditional TA group expressed the need for the training exercises to have clearer links to golf. Again, this may link to the need for familiarisation within the context of the task and the learning of TA may be easier for participants if learned within the context of golf, in this instance. While studies exclusively using traditional TA training instructions [20,32] have captured valuable verbalization data, the instruction clarity and qualitative social validation data gleaned in this present study appear to suggest the richness of verbalizations and confidence of participants could have been enhanced by including task-specific training instructions. The instruction clarity and qualitative social validation data, therefore, serve to support previous studies which have used a combination of traditional and task-specific TA training approaches [17,26,44].

Although this study has successfully investigated the training methods of TA, it is important to recognize the limitation of the lack of female representation within the study sample. Indeed, close inspection of the literature reveals that very few studies [27,29,31,45] have used female participants. Subsequently, future research is warranted to examine TA protocols by using female samples to better understand how it can be best trained and utilized.

Future implications

The overall findings from this present study appear to indicate no differences in verbalization frequency and perceptions of training effectiveness between the traditional TA training protocols outlined by Ericsson and Simon [1] and the task-specific TA training protocols designed for this study. This finding is pleasing to see as it supports the adequacy of existing methods used to train TA in the literature, both in sport and exercise psychology and beyond. Subsequently, the findings of this present study provide researchers and practitioners with the confidence in how to effectively train TA [46-48].

However, the findings from this study also appear to suggest that traditional TA training protocols can be enhanced by the use of task-specific training exercises. In an article outlining the utility of TA, Eccles, and Arsal [18] advocate the use of warm-up exercises to ensure that participants are familiar with verbalizing their thoughts out loud. Indeed, Eccles and Arsal [18] outlined the common pitfalls with applying the TA method, namely, allowing and encouraging descriptions and explanations of thoughts, no warm-up exercises, thinking aloud for too long, and concerns about reactivity. Given the meticulous consideration of training protocols used in this present study, researchers and practitioners using TA are strongly encouraged to harness Ericsson and Simon's [1] guidelines to train TA but also integrate task-specific training exercises. Indeed, previous research [17,26,28] has used a combination of traditional and task-specific instructions to train TA yet this is the first study to provide empirical evidence to support its implementation. Using a combination of existing guidelines and the recommendations gleaned from this study may enable researchers and practitioners to make reliable comparisons to previous research whilst reaping the apparent rewards of using task-specific TA training exercises. Upon analyzing the qualitative data gleaned from this study, it is clear that participants value the use of feedback and reiteration of principles when learning how to effectively TA. To the author's knowledge, this study is the first to harness social validation methods to examine participant perceptions of TA and more specifically, how to best train TA. Although it was not possible to provide bespoke feedback to participants in this present study as it would have compromised experimental control, researchers and practitioners are encouraged to monitor TA training progress (e.g., by using social validation methods such as TA checklists, measures of TA efficacy and open questioning) to ensure all participants learn how to effectively TA before data collection commences [49-51].

It is important to note that previous research has used methods to monitor the learning of TA within training protocols [26], yet similar to the TA training instructions presented in the literature, the use of such learning monitoring methods has been inconsistent. Implementing more thorough TA training procedures will not only enhance the participant's confidence of thinking aloud, but will also enhance the rigor underpinning verbalizations, and in turn, the authenticity of verbalizations captured [52-54]. Overall, the findings from this study show no differences in TA content or the perceptions of training effectiveness between the traditional TA training group and the task-specific TA training group. Notwithstanding, the findings provide some support for the contention that task-specific TA training exercises may provide some advantages over and above those offered by traditional training exercises as outlined by Ericsson and Simon [1]. We hope this study provides supplementary guidance regarding how to effectively train TA, which in turn, promotes researchers and practitioners to fully harness TA protocols as a valid means to capture in-event cognitions [55].

REFERENCES

[1] Ericsson, K.A., et al., 1993. Verbal reports as data. Cambridge: MIT Press

[2] Bloom, B.S., et al., 1950. Problem-solving processes of college students. Supplementary Educational Monographs, pp. 109.

[3] Taylor, S., et al., 2017. Usability and acceptability of an electronic pain monitoring system for advanced cancer: a think aloud study. *BMJ Support Palliat Care*, 6.

[4] McRobert, A.P., et l., 2013. Contextual information influences diagnosis accuracy and decision making in simulated emergency medicine emergencies. *BMJ Qual Saf*, 22, pp. 478-484.

[5] Banning, M. 2008. A review of clinical decision making: Models and current research. *J Clin Nurs*, 17, pp. 187-195.

[6] Ellis, A.K. 2013. Teaching, learning, and assessment together: Reflecting assessment for elementary 577 classrooms. Routledge; New York, London.

[7] Ward, P., et al., 2003. Underlying mechanisms of perceptual-cognitive expertise in soccer. *J Sport Exerc Psychol*, 25, pp. 136.

[8] Whitehead, A. E., et al., 2016. Evidence for skill level differences in the 647 thought processes of golfers during high and low-pressure situations. *Front Psychol*, 6, pp. 1974.

[9] Welsh, J.C., 2018. Thinking Aloud: An exploration of cognitions in professional snooker. *Psychol Sport Exerc*, 36, pp. 197-208.

[10] Folkman, S., et al., 2004. Coping: Pitfalls and promise. *Annu Rev Clin Psychol*, 55, pp. 745-774.

[11] Ptacek, J.T, et al., 1994. Limited correspondence between daily coping reports and retrospective coping recall. *Psychol Assess*, 6, pp. 41.

[12] Smith, R.E., et al., 1999. Can people remember how they coped? Factors associated with discordance between same-day and retrospective reports. *J Pers Soc Psychol*, 76, pp. 1050.

[13] Stone, A., et al., 1998. A comparison of coping assessed by ecological momentary analysis and 630 retrospective recall. *J Pers Soc Psychol*, 74, pp. 1670-1680.

[14] Williams, M.A., et al., 2011. Perceptual-cognitive expertise in sport and its acquisition: implications for applied cognitive psychology. *Appl Cogn Psychol*, 25, pp. 432-442.

[15] De Groot, A.D. 1978. Thought and choice in chess (Revised translation of De Groot, 1946). The Hague: Mouton Publishers.

[16] Connors, M.H., et al., 2011. Expertise in complex decision making: The role of search in chess 70 years after de groot. *Cognitive Science*, 35, pp. 1567-1579.

[17] Whitehead, A. E., et al., 2015. Examination of the suitability of collecting in event cognitive processes using Think Aloud protocol in golf. *Front Psychol*, 6, pp. 645.

[18] Eccles, D.W., 2017. The think aloud method: what is it and how do I use it? *Qual Res in Sport Exerc Hea*, 9, pp. 514-531.

[19] Schmidt, R.A., et al., 2000. Motor learning and performance: A problem-based learning approach. Champaign. IL: Human Kinetics, 58, pp. 234-235

[20] Nicholls, A.R., et al., 2008. Think-aloud: Acute stress and coping strategies during golf performances. *Anxiety Stress Coping*. 21, pp. 283-94.

[21] Van Someren, M.W., et al., 1994. Think Aloud method. A practical guide to modeling cognitive processes. London: Academic Press. pp. 1-205

[22] Elliot, D., 2010. Goal-directed aiming: two components but multiple processes. *Psychol Bull*, pp. 1-22.
[23] Lord, R. et al., 1991. Leadership and information processing: Linking perception and performance. Boston, MA: Unwin Hyman.
[24] Anderson, J.R. 1990. The adaptive character of thought. Hillsdale, NJ.
[25] Kahneman, D. 1973. Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.
[26] North, J.S., et al., 2011. Mechanisms underlying skilled anticipation and recognition in a dynamic and temporally constrained domain. *Memory*, 19, pp. 155-168.
[27] Arsal, G., et al. 2016. Cognitive mediation of putting: Use of a think-aloud measure and implications for studies of golf-putting in the laboratory. *J Sports Exerc Psychol*, 27, pp. 18-27.
[28] Runswick, O.R., et al., 2018. The effects of anxiety and situation-specific context on perceptual-motor skill: A multi-level investigation. *Psychol Res*, 82, pp. 708-719.
[29] Calmeiro, L., et al., 2011. Concurrent verbal protocol analysis in sport: Illustration of thought processes during a golf-putting task. *J Clin Sports Psychol*, 5, pp. 223- 236.
[30] Whitehead, A.E., et al., 2017. Changes in cognition over a 16.1 km cycling time trial using Think Aloud protocol: Preliminary evidence. *Int J Sports Exerc Psychol*, 15, pp. 1-9.
[31] Whitehead, A.E., et al., 2018. Investigating the relationship between cognition, pacing strategies and performance in 16.1 km cycling time trials using a think-aloud protocol. *Psychol Sports Exerc*, 34, pp. 95-109.
[32] Samson, A., et al., 2015. Think-aloud: An examination of 616 distance runners' thought processes. *Int J Sports Exerc Psychol*, 15, pp. 176-189.
[33] Mellalieu, S.D., et al., 2009. The effects of a motivational general-arousal imagery intervention upon pre-performance symptoms in male rugby union players. *J Sport Exerc Psychol*, 10, pp. 175-185.
[34] Thelwell, R.C. et al., 2006. Using psychological skills training to develop soccer performance. *J Appl Sport Psychol*, 18, pp. 254-270.
[35] Baumeister, R.F.,1986. A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *Eur J Soc Psychol*, 16(4), pp. 361-383.
[36] Arsal, G. 2013 investigating skilled and less-skilled golfers' psychological preparation strategies: The use of a think-aloud cognitive process-tracing measure.
[37] Bandura, A. 1997. Self-efficacy: The exercise of control. Macmillan.
[38] Newton, P., et al., 2008. Exploring types of educational action research: Implications for research validity. *Int J Qual Methods*, 7, pp. 18-30.
[39] Page, J., et al., 2013. The value of social validation in single-case methods in sport and exercise psychology. *J Appl Sport Psychol*, 25, pp. 61-71.
[40] MacPhail, C., et al., 2016. Process guidelines for establishing intercoder reliability in qualitative studies. *Qualitative Research*, 16, pp. 198-212.
[41] Cohen, J. 1994. The earth is round (p<0.05). *Am Psychol*, 49, pp. 997-1003.
[42] Scanlan, T.K., 1989. An in-depth study of former elite figure skaters: II. Sources of enjoyment. *J Sport Exerc Psychol*, 11, pp. 65-83.
[43] Smith, B., et al., 2017. Developing rigor in qualitative research: Problems and opportunities within sport and exercise psychology. *Int Rev Sport Exerc Psychol*, 11, pp. 101-121.
[44] Runswick, O.R., et al., 2018. The effects of anxiety and situation-specific context on perceptual-motor skill: A multi-level investigation. *Psychol Res*, 82, pp. 708-719.
[45] Kaiseler, M., et al., 2013. Gender differences in stress, appraisal, and coping during golf putting. *Int J Sport Exerc Psychol*, 11, pp. 258-272.
[46] Adams, D., et al., 2016. Developing specialized youth soccer coaching qualifications: An exploratory study of course content and delivery mechanisms. *International Sport Coaching Journal*, 3, pp. 31-45.
[47] Balagué, N., et al. 2015. Intentional thought dynamics during exercise performed until volitional exhaustion. *J Sports Sci*, 33, pp. 48-57.
[48] Cumming, J., et al. 2004. The Influence of an Imagery Workshop on Athletes' 570 Use of Imagery. Athletic Insight, 6, pp. 52-73.
[49] Hsieh, H. F., et al. 2005. Three approaches to qualitative content analysis. *Qual Health Res*, 15, pp. 1277-1288.
[50] O'Brien, M., et al. 2009. Goal-setting effects in elite and nonelite boxers. *J Appl Sport Psychol*, 21, pp. 293-306.

[51] Ward, P., et al. 2003. Underlying mechanisms of perceptual-cognitive expertise in soccer. *J Sport Exerc Psychol*, 25, pp. 136.

[52] Welford, A.T 1967. Single-channel operation in the brain. *Acta Psychologica*, 27, pp. 5-22.

[53] Ryeong, Ha.O., et al. 2018. Food advertising literacy training reduces the importance of taste in children's food decision-making: A pilot study. *Front Psychol*, 27, pp. 1293.

[54] Güss, C.D., 2018. What is going through your mind? Thinking aloud as a method in cross-cultural psychology. *Front Psychol*, 13, pp. 1292.

[55] Zeijlmans, K., et al. 2019. Heuristic decision-making in foster care matching: Evidence from a think-aloud study. *Child Abuse Negl*, 88, pp. 400-411.