



Journal of Sports Sciences

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/rjsp20>

Developing the Coach Analysis and Intervention System (CAIS): Establishing validity and reliability of a computerised systematic observation instrument

Christopher Cushion^a, Stephen Harvey^b, Bob Muir^c & Lee Nelson^d

^a School of Sport Exercise and Health Sciences, Loughborough University, Loughborough, UK

^b Department of Physical Education & Sport Studies, Faculty of Education, Sport & Tourism, Institute of Sport & Physical Activity Research, University of Bedfordshire, Bedford, UK

^c Carnegie Faculty of Sport, Leeds Metropolitan University, Leeds, UK

^d Department of Sport Sciences, University of Hull, Hull, UK

Published online: 05 Dec 2011.

To cite this article: Christopher Cushion, Stephen Harvey, Bob Muir & Lee Nelson (2012) Developing the Coach Analysis and Intervention System (CAIS): Establishing validity and reliability of a computerised systematic observation instrument, Journal of Sports Sciences, 30:2, 201-216, DOI: [10.1080/02640414.2011.635310](https://doi.org/10.1080/02640414.2011.635310)

To link to this article: <http://dx.doi.org/10.1080/02640414.2011.635310>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Developing the Coach Analysis and Intervention System (CAIS): Establishing validity and reliability of a computerised systematic observation instrument

CHRISTOPHER CUSHION¹, STEPHEN HARVEY², BOB MUIR³, & LEE NELSON⁴

¹*School of Sport Exercise and Health Sciences, Loughborough University, Loughborough, UK,* ²*Department of Physical Education & Sport Studies, Faculty of Education, Sport & Tourism, Institute of Sport & Physical Activity Research, University of Bedfordshire, Bedford, UK,* ³*Carnegie Faculty of Sport, Leeds Metropolitan University, Leeds, UK,* and ⁴*Department of Sport Sciences, University of Hull, Hull, UK*

(Accepted 22 October 2011)

Abstract

We outline the evolution of a computerised systematic observation tool and describe the process for establishing the validity and reliability of this new instrument. The Coach Analysis and Interventions System (CAIS) has 23 primary behaviours related to physical behaviour, feedback/reinforcement, instruction, verbal/non-verbal, questioning and management. The instrument also analyses secondary coach behaviour related to performance states, recipient, timing, content and questioning/silence. The CAIS is a multi-dimensional and multi-level mechanism able to provide detailed and contextualised data about specific coaching behaviours occurring in complex and nuanced coaching interventions and environments that can be applied to both practice sessions and competition.

Keywords: *Coach behaviour, coach observation, behaviour analysis*

Introduction

A significant focus over that last 30 years has been to systematically identify the behaviour of coaches and teachers using descriptive-analytical systems (Gilbert & Trudel, 2004; Mesquita, Rosado, Januario, & Barroja, 2008; Smith & Cushion, 2006). Since the pioneering work of Tharp and Gallimore with legendary coach John Wooden (Tharp & Gallimore, 1976), a wide range of literature has evolved describing coaches' behaviour in training and competition, with observation systems being used to compare coaches behaviour across a diverse range of sports such as soccer (e.g., Cushion & Jones, 2001; Ford, Yates, & Williams, 2010; Potrac, Jones, & Cushion, 2007), basketball (e.g., Bloom, Crumpton, & Anderson, 1999), golf (e.g., Schempp et al., 2004), baseball (e.g., DeMarco, Mancini, & Wuest, 1996) ice hockey (e.g., Trudel, Côté, & Bernard, 1996), and tennis (e.g., Lacy & Goldston, 1990). This work has gone some way to creating a 'data-base' of coaching behaviour in a range of sports, the constant production of articles over this period suggesting this

methodology is well established for analysing coaching (Cushion, 2010; Smith & Cushion, 2006).

While systematic observation has become a recognised methodology, and observation of coaches has resulted in a wealth of information about coach behaviour, observation instruments are limited in that they only measure direct styles of coaching that are often stripped of context (Kahan, 1999; Sherman & Hasan, 1984; Smith & Cushion, 2006). Behavioural research is not able to provide insight into the contextual factors or cognitive process underlying these behaviours (Gilbert & Trudel, 2004). Indeed, behavioural assessment has been criticised for being too simplistic (Strean, 1995) and limited in its contribution to coach effectiveness (Bloom, Durand-Bush, & Salmela, 1997). Therefore, there have been calls to extend overly simplistic observation systems (Smith & Cushion, 2006), and provide more sensitive evaluation methods (Gilbert & Trudel, 1999; Brewer & Jones, 2002).

Coaching is still a developing field and by its nature requires descriptive research to develop an understanding and accumulation of knowledge about what

coaches actually do (Gilbert & Trudel, 2004; Potrac et al., 2007). Systematic observation provides descriptive baseline data of 'actual' (demonstrated in context) coaching behaviour and is the precursor for the use of accompanying methodologies and higher-level research (Brewer & Jones, 2002; Gilbert & Trudel, 2004; Potrac et al., 2007). Systematic observation has an essential functional step in answering the question: What is going on here? (Brewer & Jones, 2002). Moreover, the breadth of coaching domains (e.g. participation, development, elite) means that despite the growing body of work undertaken, coaching remains under-researched.

A number of systems currently exist for observing athletic roles or recording behaviour in sports or physical education (Darst, Zakrajsek, & Mancini, 1989), with probably the best known and most widely used in coaching being the Coach Behaviour Assessment System (CBAS) (Smith, Smoll, & Hunt, 1977) and the Arizona State University Observation Instrument (ASUOI) (Lacy & Darst, 1984). Arguably, these existing systematic observation instruments report a less nuanced picture of coach-athlete interaction because they are not sufficiently sensitive to context dependent coaching practice, and are by definition limited to the scope of the observation tool (Cushion, 2010). The accuracy of any behavioural records emergent from observational research relies upon the ability of the instrument to distinguish and measure crucial coaching behaviours (Brewer & Jones, 2002; Horn, 2002). In an examination of expert and non-expert tennis coaches, Abraham, Collins, Smethurst, and Collins (1997) compared the Coach Behaviour Assessment System (CBAS) to an extended observation system. While the extended system found two significant differences in behaviour, the less sensitive CBAS found no significant differences. Moreover, Cushion and Jones (2001) note the frequency of behaviours such as 'humour' and 'uncodable' behaviour with the ASUOI, while Smith and Cushion (2006) subdivided the ASUOI category of silence to give a more detailed perspective of that behaviour. Researchers (e.g., Bloom et al., 1999; Brewer & Jones, 2002; Ford et al., 2010; Smith & Cushion, 2006) have modified and extended existing instruments to increase their breadth and sensitivity.

Clearly, existing systematic observation instruments while generally useful have not been designed with context in mind, assuming coaching is homogenous across contexts. This factor may, in turn, lead to overly simplistic description and therefore discrepancies between the situational behaviours and those actually recorded (Brewer & Jones, 2002). The structure and content of coaching will be different across sports, while within sports coaches deal with differing situations of competition, organisation and training (Côté, Salmela, Trudel, Baria, & Russell

1995; Lyle & Cushion, 2010). Furthermore, within sessions different instructional strategies will be employed to adjust the task cognitive load for different practice levels (Hodges & Franks, 2002; Pereira, Mesquita, & Graça, 2009). Existing systems are not grasping this subtlety because the behaviours defined within the instruments simply do not allow for these differences.

Frequently, research using new (e.g., Periera et al., 2009) or amended instruments use technology to record coaching behaviour, but still utilise a 'paper and pencil' notational methodology for analysis. However, given the rapid developments in computer technology, it is surprising that only one computerised system has been reported in the literature specifically to analyse coaching, the computer-aided coaching analysis system CAI (Franks, Johnson, & Sinclair, 1988) and two for analysing physical education lessons PETEACH (Hawkins, Sharpe, & Ray, 1994) and ALT-PE Microcomputer Data Collection System (Metzler, 1989). Computer technology seems relatively untapped as a resource to help develop systematic observation. For example, technology helps with data collection and analysis and allows for the summary and display of both quantitative and qualitative 'video' data (More, 2008). These steps alone offer improved opportunities for interventions, with the timely return of meaningful analysis and feedback presented in 'video' format. Coaches can observe themselves coaching, rather than trying to grasp a graphical representation of their actions and behaviour (Stratton, Reilly, Williams, & Richardson, 2004).

Moreover, any 'paper and pencil' methodology, regardless of the sophistication of the system can only ever record coaching behaviour in a uni-dimensional and sequential manner (i.e. one behaviour must follow another). The nature of coaching in practice sessions and competition however is complex and dynamic in ill-structured changing environments (Cushion, 2007; Nash & Collins, 2006). Coaches do more than one thing at once. Therefore, a methodology used within any particular setting should be able to deal with these levels of complexity and provide as great a quantity of valid and reliable data as possible: any instrument unable to identify behaviours important within specific coaching contexts limits the functional information that can be gained (Brewer & Jones, 2002). Even within a particular sport at specific competitive levels, coach behaviour could be examined in both competition and practice/training environments (which itself has different states) and at different times in the season. Therefore, the development of a more sophisticated analysis tool, coupled with computer technology enables a hierarchical or multi-level form of event recording; thus capturing the complexity of coaching behaviour and describing the subtlety and nuances of

coaching practice more readily. The application of technological innovation provides the opportunity to breathe new life into systematic observation methodology, with Stratton et al. (2004) suggesting that analysis software offers “a number of exciting opportunities for providing feedback in innovative and creative ways” (p.132).

To expand and extend coach behaviour research, digital video and computerised analysis software has been used to develop the Coach Analysis and Intervention System (CAIS). The CAIS is a systematic observation instrument that currently uses ‘Sports-Code’TM analysis software as a template, thus creating a computerised analysis system with a hierarchical or multi-level form of event recording. This process enables a trained observer to capture up to five levels of simultaneous behaviour (primary behaviour plus secondary detail, context, recipient, timing, content), in training/practice or match/competition settings, where existing observation instruments typically have one or two levels (primary behaviour plus recipient OR timing). This depth of analysis means that the data provided by the CAIS provides a more complete description of the coaches’ behaviour than any other current observation system and thus equips both coach and researchers with detailed, nuanced and explicit information to examine and reflect upon, and act as a basis for intervention. Because it is based on digital video recordings of coaching practice the systems outputs can provide powerful video images or ‘clips’ of any selected behaviour, combinations of behaviours, or coaching ‘moments’.

Therefore, the broad aim of this research was to develop and validate a computerised, multidimensional, hierarchical analysis system able to identify specific behaviours that occur within complex coaching environments. The specific purpose of this paper is to describe the system’s evolution and the process for establishing the validity and reliability of this new systematic observation instrument.

Method: The five-stage development process

If systematic observation is to accurately describe the interactive behavioural patterns within sports coaching, and provide foundational research for the development and evolution of an understanding of coaching, it is essential that the methodology for the validation of specific observation instruments is transparent. Many behavioural categories within systematic observation instruments developed deductively are not discrete and have not been formally validated (Brewer & Jones, 2002; Cheffers, 1977). The development of the Coach Analysis and Intervention System (CAIS) is explained and the methodology described. To establish the validity and

reliability of the CAIS the five-stage development process as outlined by Brewer and Jones (2002) and recently used by Ford et al. (2010) acted as a template.

Following institutional ethics approval, the process included: (i) an initial literature review and development of the instrument, establishing initial face validity and pilot work; (ii) observer training; (iii) system use and amendments to establish content validity; (iv) establishing face validity with experts in the domain (coaches and researchers); (v) establishing intra- and inter-observer reliability.

Stage 1: Literature review and initial instrument design

The first step was a review of literature examining behavioural categories of existing observation instruments. These included: Coach Observational System (Quarterman, 1980); Coaching Behaviour Recording Form (Langsdorf, 1979); Coach/Athlete Observation Schedule (Rushall, 1977); Lombardo Coaching Behaviour Analysis System (Lombardo, Faraone, & Pothier, 1983); Feedback Analysis Profile (Sinclair, 1989); Arizona State University Observation System (Lacey & Darst, 1984) Coaching Behaviours Observational Recording System (Tannehill & Burton, 1989); Coach Behaviour Analysis System (Smith et al., 1977); Systematic Analysis of Pedagogical Content Interventions (Gilbert, Trudel, Gaumont, Larocque, 1999). No single existing instrument was sufficiently encompassing for development into a new computerised system therefore behaviour definitions were developed specifically for the CAIS.

The expertise of the research team enhanced the face validity of the initial version of the CAIS, with a combined total of 50 years experience as coaches, all certified as coaches, and published extensively on coaching and coach behaviour. In addition, and prior to a pilot test, five expert coaches (certified to the highest level in their sport; 10–30 years experience, employed full-time as coaches) were asked to evaluate the definitions, to determine clarity of the wording and whether the categories captured coaching practice and were given questions adapted from Safrit (1986); Brewer and Jones (2002) and Periera et al. (2009): (i) Are there elements omitted from the behaviour categories? (ii) Are any behaviours erroneously included? (iii) Does the content reflect coaching behaviours in practice and game settings? Only minor revisions were made to the categories/definitions as a result of this testing of face validity of this version.

This version of the instrument was piloted with an international youth soccer coach. At a week-long pre-tournament training camp, the coach was filmed over six sessions which yielded nine hours of coaching behaviour footage. Two experienced coders were

given the computerised instrument and asked to code sessions, reviewing both coding decisions, and behavioural definitions and categories. Using the software helped to identify relationships between categories, with the 'layering' of coach behaviour, and in identifying multiple behaviours in single coaching moments. This process also identified 'non-logical' relationships, for example, multiple primary (e.g. feedback and silence) or secondary behaviours (e.g. recipient – individual, group, team or other) occurring simultaneously that are in fact contradictory. The outcome of Stage 1 was a working version of the Coach Analysis Intervention System (CAIS).

Stage 2: Observer training

Brewer and Jones (2002) argue that, "a rigorous observer-training program is of particular importance in determining the reliability and objectivity of a systematic observation instrument" (p.143). The initial CAIS was used to train the coders ($n = 3$) in computerised systematic observation. The instrument had undergone preliminary pilot testing (Stage 1 above) and gave a more in-depth description of the coaches' behaviour using both primary behaviour and secondary detail of behaviour (e.g., feedback could be considered as corrective, general or specific and according to recipient, its content, and when it occurred). Thus the CAIS accounted for all coaching behaviour, significantly reducing the use of the 'uncodable' category. When a primary behaviour was recognised, secondary detail of behaviour could be classified. Figure 1 reflects the hierarchical nature of the recording process developed and identifies the 'steps' involved. In step one, coders would first choose one of the 23 'primary' behaviours. In steps two to six the coders would then record the secondary detail of the behaviours: Step two, the nature of the activity (i.e. match/game or training/practice state – see Table I), step three, the recipient (e.g., team, individual), step four, timing (pre, concurrent, post), step five, content (technical, tactical or other), and if required step 6, the nature of the questioning (convergent/divergent) and/or silence (on/off task). For example:

Step 1. *What was the behaviour?* Behaviour = 'Instruction' (i.e. 'Johnny, mark their number 10')

Step 2. *During which activity?* (match/game or practice/training state category) Activity = 'Match/Game state – Performance 1st Quarter'

Step 3. *Who was the recipient?* Recipient = 'Individual'

Step 4. *When was it given (timing?)* Timing = 'Post' (i.e. it was a response to 10 receiving the ball)

Step 5. *What was the content?* Content = 'Tactical' (due to the relation of the coach's statement to

cognitive strategies involving player decision-making, rather than technical skill such as passing dribbling or shooting- see Table I)

Step 6. *What was the nature of questioning and/or silence? (IF REQUIRED)* Not applicable in this instance

These descriptions and process helped train the coders to recognise the nuanced nature of coaches' intervention behaviour, rather than leaving description at a more simplistic and superficial level. A five phase process for observer training outlined by van der Mars (1989, pp. 75–77) was utilised. This included: a) orientation to the system; b) learning the categories; c) using the 'coding timeline'; d) initial coding practice; e) live observation practice. Coders were trained over a four-week period. First, coders were introduced to the basic purposes of the system including the definitions and descriptions of the CAIS behaviours. Second, while stressing the need for 'tunnel vision' (van der Mars, 1989 p. 75 original emphasis), all behavioural classifications (cf. Darst et al., 1989) were learned using the SportsCodeTM software. Third, coders learned to use the 'coding timeline' correctly within the SportsCodeTM software following the six coding steps previously outlined. In phase four, videotaped excerpts of training sessions and games were coded. Initially, only ten-minute sections of footage were coded and coders were asked to classify certain behaviours (e.g., feedback, management, instruction behaviours) so that these could be reviewed, and any misclassifications discussed before coding a further ten-minute section of data. More behaviours were introduced as the training developed and the coders were consistent in classifying these behaviours. Modelling and verbal/non-verbal behaviours were added, (e.g., humour, praise, silences) before the various secondary detail of behaviours (i.e., recipient, timing, content) were introduced. Van der Mars (1989) highlights that as the amount of behavioural classifications increases, reliability decreases, as a result with increasing system complexity, the threshold point for reliability scores should decrease. The phased training of the coders ensured that reliability was maintained as the number, and complexity of behavioural classifications increased. Finally, one full training session/ game with the content of the whole behavioural classification, for both primary and secondary behaviours was coded.

Building on the Stage 1 pilot, data from three University sports coaches, coaching three different sports were used to aid the development and validation of the CAIS. In total 26 training sessions ($n = 11$ hockey; $n = 8$ basketball; $n = 7$ volleyball) and 22 games ($n = 7$ hockey; $n = 7$ basketball; $n = 8$ volleyball) were filmed over the length of one season/year. All coded behaviours were saved on a

Steps in recording & Primary/Secondary Behaviours and/or Category†	^STEP 1 Primary Behaviour*	STEP 2 Match/ Practice State	STEP 3 Recipient	STEP 4 Timing	STEP 5 Content	STEP 6 Questioning	STEP 6 Silence
Physical Behaviours	1. Positive Modelling	√	√	√	√	x	x
	2. Negative Modelling	√	√	√	√	x	x
	3. Physical Assistance	√	√	√	√	x	x
Feedback	4. Specific Feedback Positive	√	√	√	√	x	x
	5. Specific Feedback Negative	√	√	√	√	x	x
	6. General Feedback Positive	√	√	√	√	x	x
	7. General Feedback Negative	√	√	√	√	x	x
	8. Corrective Feedback	√	√	√	√	x	x
Instruction	9. Instruction	√	√	√	√	x	x
Verbal/Non-verbal	10. Humour	√	√	√	x	x	x
	11. Hustle	√	√	√	x	x	x
	12. Praise	√	√	√	x	x	x
	13. Punishment	√	√	√	x	x	x
	14. Scold	√	√	√	x	x	x
	15. Uncodable	√	√	√	x	x	x
Silence	16. Silence	√	x	x	x	x	√
Questioning	17. Question	√	√	√	x	√	x
	18. Response to Question	√	x	x	x	x	x
Management	19. Management - Direct	√	√	x	x	x	x
	20. Management – Indirect	√	√	x	x	x	x
	21. Management – Criticisms	√	√	x	x	x	x
Others	22. Confer with Assistant	√	x	x	x	x	x
	23. Verbal Protocol Analysis	√	x	x	x	x	x

Figure 1. Figure outlining the relationships between the steps involved in the recording of primary and secondary behaviours in the CAIS. Notes: *Behavioural definitions are provided in Table I. †Actual names of secondary detail of behaviour and their definitions are provided in Table I. ^The 'steps' in the coding process are explained in the text of the paper. √/- refers to secondary detail of behaviour requiring coding in steps 2–6 of the coding process (see text); x – refers to secondary behaviour not requiring coding in steps 2–6 of the coding process (see text).

Table I. Training/game states CAIS comparison to Ford et al. 2010.

Behavioural Classification (CAIS) Practice/Training States	Behavioural Description	Behavioural Classification (Ford et al.) Training Form	Behavioural Description
Physiological	Primary goal of the physiological aspects of the game (e.g. warm-up, cool-down, conditioning, stretching).	Fitness	Improving fitness aspects of the game without the ball.
Technical Practice	Individual or group activity covering isolated technical skills under limited or no pressure.	Technical	Isolated technical skills unopposed alone or in a group.
Skills Practice	Individual or with a group covering technical skills under pressure from opponent(s).	Skills	Re-enacting isolated simulated game incidents with or without focus on particular technical skill.
Functional Practice	Re-enacting isolated simulated game incidents with or without a focus on a particular skill. These can be unopposed (i.e. under no pressure from opponents(s) such as shadow play/game simulated patterns of play/movement) or opposed.	Not Defined	
Playing States Phase of Play	Attack vs. defence play covering team cognitive strategies used to outsmart opponents which involve ONLY one team of players scoring, or the two teams of players scoring in different ways (i.e. attackers score in a main regulation goal/basket and the other team runs the ball out through some marker cones/over a line, or by outletting the ball to a feeder/coach etc.).	Playing Form Phase of Play	Uni-directional match-play towards one goal.
Possession Game	No goals/targets in which retention of possession rather than scoring is the primary objective.	Not defined	(Incorporated in conditioned game)
Conditioned Game	Restrictions/variations to rules, goals, or area of play, but with both teams scoring in the same way.	Conditioned game	As small-sided games but with variations to rules, goals or areas of play.
Small Sided Game	Two goals, realistic to regulation rules, with both teams scoring in the same way.	Small Sided game	Match play with reduced number of players and two goals.
Full Sided Game	Two goals, regulation rules and players (i.e. 11 vs. 11 in case of field hockey), with both teams scoring in the same way.	Not defined	
Management/ Transition States State – Other*	Management/Transition States Time when coaches are managing/addressing the team of players to explain up-and-coming practices, transition or move players from one practice state (i.e. technical practice) to a NEW or DIFFERENT practice state (i.e. skills practice), and time when players are taking a breaks from practice (i.e. water breaks).	Not defined	
Match/Game States Performance States within Match Competition	Performance (includes 1 st Half/Q ^{ter} /Game, 2 nd Half/Q ^{ter} /Game, 3 rd Q ^{ter} /Game, 4 th Q ^{ter} /Game) Coaching behaviours that occur during performance (i.e. re-enforcement of good play, verbal or non-verbal scolds etc.).	Not Defined	
Non-Performance States within the Match competition			
Timeout	Coaching behaviours that occur during a timeout (i.e. instruction, re-enforcement of good play, verbal or non-verbal praise, verbal or non-verbal scolds etc.).		
Half-Time	Coaching behaviours that occur during the half-time break in hockey and basketball (i.e. instruction, re-enforcement of good play, verbal or non-verbal praise, verbal or non-verbal scolds etc.).		
End of Quarter/Game	Coaching behaviours that occur during an end of a quarter in basketball or the conclusion of one of the games in volleyball (i.e. instruction, re-enforcement of good play, verbal or non-verbal praise, verbal or non-verbal scolds etc.).		

(continued)

Table I. (Continued).

Behavioural Classification (CAIS) Practice/Training States	Behavioural Description	Behavioural Classification (Ford et al.) Training Form	Behavioural Description
Non-Performance States outside of the formal Match competition	Non-Performance States outside of the formal Match competition	Not Defined	
Pre-Match	Coaching behaviours that occur before performance commences (i.e. during the pre-match team talk etc.).		
End of Match	Coaching behaviours that occur at the conclusion of the whole match/competition (i.e. instruction, feedback of good play, verbal or non-verbal praise, verbal or non-verbal scolds etc.).		

'*timeline*' on the hard drive of the computer for later analysis. Coaches also completed post practice/competition structured interviews, an interim and final semi-structured interview, based on the findings of the research, and three stimulated recall sessions for both practice and competition.

Movie-test development procedure. Intra- and Inter-Observer Reliability (IOR) was established using test-retest reliability on a movie-test, and this procedure also determined and checked systematic observation 'competence'. A movie-test was used, and defined in this case as a series of coaching including 45 clips that had at least one example of each behaviour from the CAIS, and not more than three. The movie-tests were comprehensive covering all of the primary and secondary detail of behaviours in the CAIS, in both training and competition contexts. 'Real' coaching footage was coded, but testing in this way ensured that all behaviours occurred. To record behaviours from different coaches, the coaching behaviours chosen for the movie-test were from three different team sport activities and included footage from coaching sessions not included in the validation study. This process gave the coders the opportunity to gain reference to behaviours from within a specified coaching context. During the test, the coaching intervention behaviours were recorded on a pre-prepared coaching timeline linked to a video of the various instances of coaching intervention behaviour from the CAIS. This process necessitated the coders following steps 1–6 of the coding process outlined in Figure 1 and examined coder competence to code the data from the CAIS.

Movie-test implementation procedure. The movie-tests were organised to ensure that recording of the behaviours seen by each coder was independent, and so there was no possibility of visual or oral interaction between the coders during the test. Each clip was played in-turn allowing 10 seconds to code before it was replayed for a second time, with another 10

seconds to code before the next clip. Primary and secondary detail of behaviour were recorded for each clip on a pre-prepared coding timeline. A one-week gap between tests is recommended to allow memory lapse to occur (Darst et al., 1989). One week later the *test-retest* movie-test was repeated with the exact procedure outlined above. To control for any remembering of the clips from the previous test, the clips were randomised and presented in a different order. An IOR percentage of 0.80 was set for all reliability coding (Gilbert et al., 1999; van der Mars, 1989).

Results. All coders ($n=3$) successfully completed the movie-test, and the re-test movie-test, which included both primary and secondary detail of behaviours. Scores for primary behaviours ranged from 0.80 to 0.85 and secondary detail of behaviours ranged from 0.81 to 0.85.

Stage 3: Amending an existing systematic observation instrument and content validity

Training using the existing CAIS with season-long data allowed the instrument to be operationalised and tested 'in-vivo'. Using the instrument in this way allowed feedback about its utility and the accuracy of definitions, whether behaviours needed to be added, or even if behaviours were erroneous. This process identified some ambiguity in the definitions (e.g., the relationships of praise and general reinforcement (+); as well as the relationships between general negative reinforcement (–), scolds and management criticisms). In addition, some of the existing behaviours were limited in describing some of the coaches' intervention behaviours. The outcome of this for example, was to add an 'instruction' category to the instrument, as certain behaviours that were instructional could not be coded under any other category. Similarly, additional descriptors to questioning (i.e., convergent and divergent) and also to silence (i.e., on or off task) were added.

One of the final major additions to the existing CAIS was the consideration of practice or training 'states' and game/match 'states'. A recent and on-going development in systematic observation research has been to consider coach behaviour according to practice activity. Ford et al.'s (2010) definitions for practice or training 'states' were utilised (see Table I) but amended and developed for greater clarity in coding. For example, 'functional practices' were added and separated into 'opposed' and 'unopposed' formats.

Furthermore, Ford et al. (2010) omitted time in which coaches may have been engaged in management or transitions from one practice/training state to another. Thus, a 'state – other' category and definition was added. Matches/games 'states' were developed that included separating the first, second, third and fourth halves/quarters/ and states such as timeout, half-time, pre-match and end of match were added. The addition of this level of analysis not only enabled the recording of the duration of the practice/training and/or match/game 'states' but also located the coaches intervention behaviour within these 'states', thus offering contextualised description of intervention behaviours. The amended CAIS was used to hold a final coding training session with the coders to familiarise them with the adapted behavioural definitions and provide detailed examples of the 'new'/'additional' behaviours and descriptors. This also presented an opportunity to gain feedback from the coders on the amendments before proceeding onto Stage 4 of the system validation process.

Stage 4: Establishing face validity

Face validity relates to the need to show that the instrument is on target with its goals and objectives when compared to non-relevant instrumentation (Brewer & Jones, 2002; Cheffers, 1977; Cushion & Jones, 2001). Following revisions from Stages 1–3 face validity was confirmed through interaction with specialists in the relevant field of study (Brewer & Jones, 2002; Ford et al., 2010; Vogt, 1999). It was important to establish that the CAIS would capture the practice and game behaviours of coaches and that the developed categorical definitions were appropriate.

Procedure from Stage 1 was followed, but with a larger panel of 10 expert coaches and three researchers. The coaches were from a range of sports and reviewed and validated the definitions, providing written feedback about the developed categories of the CAIS. All of the coaches had a minimum of 5 years coaching experience, and coached university, National League, or international athletes. The panel review also included researchers with expertise in coaching research, and the design and use of

systematic observation instruments. Each had published work using systematic observation, and was comfortable in providing comment on the definitions developed, the integration of the system, and the completeness of the instrument (Brewer & Jones, 2002). Panel members considered the instrument (including the behavioural classifications, behavioural and state descriptors, and examples of behaviour), and a conceptual 'map' of the instrument (see Figure 1). The panellists were given the same questions from Stage 1, and their comments were reviewed by the research team looking for discrepancies and issues identified from the review questions. In general, there was agreement from the panel. Discrepancies were identified around the competition element of the instrument. These included issues of sport specific language in descriptors, and terminology around 'game phases' in different sports. For example, while the definitions of game states were deemed sound, the descriptor could vary by sport, for example, first-half, first inning, first round, first game, as an initial period of performance. Additional comments were about refining behaviours that were conceptually related (e.g., instruction and prompt, and reinforcement and feedback) so that they could be distinguishable when using the instrument. The research team, following these comments, clarified the feedback and modified the instrument if necessary and some revisions were made to the categories and/or definitions. The finalised instrument is presented in Tables II–VII and visually in Figure 1. This procedure ensured the CAIS had content and face validity and was therefore an appropriate instrument to record coach behaviour in practice and competition settings.

Stage 5: Establishing intra- and inter-observer reliability

To establish intra- and inter-observer reliability, the collected data previously described were used. A range of approximately 30% of the data for both intra- and inter-observer reliability checks was assigned (Baumgartner, Jackson, Mahar, & Rowe, 2007). Therefore, intra- and inter-observer reliability checks were completed on a total of $n = 15$ sessions ($n = 9$ training and $n = 6$ games). The process of completing both intra- and inter-reliability from the study promoted robust coder reliability, in that, there was not simply a cross sectional 'snapshot' of the coders but monitoring over a longitudinal period of coding. Thus, unlike Brewer and Jones (2002) there was control for procedural reliability and observer drift of the coders. In addition, the use of three coders strengthened the observer reliability process, as gaining reliability was more difficult than simply having two people agree, or a single person gaining intra-agreement with themselves (van der Mars, 1989).

Table II. Primary behaviours (Step 1).

Behavioural Classification	Behavioural Description
Primary Coaching Behaviour	
1. Positive Modelling	Skill demonstration- with or without verbal instruction that shows performer the correct way to perform.
2. Negative Modelling	Skill demonstration- with or without verbal instruction that shows the performer the incorrect way to perform.
3. Physical Assistance	Physically moving the performer's body to the proper position or through the correct range of movement.
4 & 5. Specific Feedback (positive or negative)	Specific verbal statements (either positive or supportive OR negative or unsupportive) <i>that specifically aim to provide information about the quality of performance</i> (can be delivered concurrently or post). e.g. <u>Positive</u> : 'Good you followed through well', 'Good drag', 'that was good defending', 'I liked the way that you got nice and low in the tackle', 'good serve', 'excellent counter attack/fast break'. <u>Negative</u> : 'Don't lose sight of the ball and your man', 'Don't force the pass', 'the attack is too slow', 'you're swinging too early', 'you've got to talk, guys', 'you maybe got caught a bit too wide'.
6 & 7. General Feedback (positive or negative)	General verbal statements OR non-verbal gestures (either positive or supportive OR negative or unsupportive (can be delivered concurrently or post). e.g. <u>Positive</u> : 'Well tried', 'well done', 'good job', 'much better', 'that's lovely', 'I like that', clapping. <u>Negative</u> : 'don't do that again', 'oh, guys, please', 'that was rubbish', 'you got that wrong', 'not from there', holding head in hands.
8. Corrective Feedback	Corrective statements that contain information that <i>specifically aim to improve the player(s) performance at the next skill attempt</i> (can be delivered concurrently or post). e.g. 'it would help if your stance was not a metre wide', 'you probably do not want to be on the same side as your team-mate initially', 'move your feet and not your stick', 'pass it earlier next time', 'force them away from the goal when they are attacking', 'we need to increase the ball pace'.
9. Instruction	Verbal cues, reminders or prompts to instruct / direct skill or play related to player(s) performance. e.g. 'Talk', 'Press', 'Keep them there', 'Mark up', 'Go with the runner', 'Move left', 'Now run left', 'Pass the ball', 'Take a shot', 'Get behind them', 'Go' (to activate or initiate a player's behaviour during a drill).
10. Humour	<i>Jokes or content designed to make players laugh or smile.</i> e.g. 'Have you got steel toe caps in those trainers?', 'Have you eaten steak for your lunch today?'
11. Hustle	Verbal statements or gestures <i>linked to effort</i> to activate or intensify previously directed behaviour. e.g. 'You can do it, keep working' or 'Pace, pace, pace', 'Go, go, go', 'Let's go'.
12. Praise	Positive or supportive verbal statements or non-verbal gestures <i>which demonstrate the coach's general satisfaction or pleasure to a player(s) that DO NOT specifically aim to improve the player(s) performance at the next skill attempt.</i> e.g. 'your work rate was excellent today', 'good effort', a smile, pat on the back or thumbs up.
13. Punishment	Specific punishment following a mistake. e.g. 'give me 20 press-ups', 'three laps', 'five sit-ups'.
14. Scold	Negative or unsupportive verbal statements or non-verbal gestures demonstrating displeasure at a player(s) <i>that DO NOT specifically aim to improve the player(s) performance at the next skill attempt.</i> e.g. Shaking of the head or arms held aloft, swearing at a player(s)
15. Uncodable	Not clearly seen or heard, not belonging to any other category.
16. Silence	Coach is silent this can be on- or off-task. (See secondary questioning behaviours below for definitions of on- and off-task).
17. Question	Coach asks a question about skill, strategy, procedure or score, the status of a player's injury, about the welfare of a player, to a match official, etc. (see secondary questioning behaviours below for specific examples).
18. Response to question	Coach responds to a question that may or may not be directly be related to practice or the match competition.
19. Management – Direct	Management that is practice/match competition related coach behaviour contributing directly to practice/ match competition or explaining how to execute the skill, drill or game. e.g. 'I need you in three groups of four' or behaviours such as target setting, refereeing or scoring in practice, or verbal explanations such as 'during this drill I want', 'the purpose of this task is to. . .', calling timeouts in matches, asking the bench to make a sub in matches, explaining the starting line-up for a match, re-organising line-ups during timeouts, etc.
20. Management – Indirect	Management that is practice related coach behaviour, not contributing directly to practice/the match competition. e.g. Collect all the balls and bibs, sending players for a water break, travel arrangements for fixtures/up-and-coming practice sessions.
21. Management – Criticisms	Management that demonstrates displeasure at the player(s) behaviour or match official's decisions. e.g. <u>Players</u> : 'You're late again!', 'leave the balls alone while I'm speaking', 'we are taking too long to get this practice organised', 'we need to make our substitutions more frequently', 'listen to the person who is speaking, please', 'girls, your attitude stinks', 'if you don't want to do this practice we can run instead', 'its fine, we will finish practice late then', 'stop whining at each other', 'stop being lazy'. <u>Officials</u> – 'that's a carry', 'that's a penalty', 'that's an unbelievable call', 'that was a foul'.
22. VPA	Coach engaged in Verbal Protocol Analysis ('think aloud techniques' verbalising their actions, communications, thoughts, and feelings).
23. Confer with Assistants	Coach confers with assistants to talk about, manage or reflect on anything concerned with the practice (i.e. management and organisation of practice, players' performance within practice and in games) which is related to current training episode.

Table III. Secondary detail of coach behaviour: Training/game states (Step 2).

Behavioural Classification	Behavioural Description
Practice/Training States	
Physiological	Primary goal of the physiological aspects of the game (e.g. warm-up, cool-down, conditioning, stretching).
Technical Practice	Individual or group activity covering isolated technical skills under limited or no pressure.
Skills Practice	Individual or with a group covering technical skills under pressure from opponent(s).
Functional Practice	Re-enacting isolated simulated game incidents with or without a focus on a particular skill. These can be unopposed (i.e. under no pressure from opponents(s) such as shadow play/game simulated patterns of play/movement) or opposed.
Playing States	
Phase of Play	Attack vs. defence play covering team cognitive strategies used to outsmart opponents which involve ONLY one team of players scoring, or the two teams of players scoring in different ways (i.e. attackers score in a main regulation goal/basket and the other team runs the ball out through some marker cones/ over a line, or by outletting the ball to a feeder/coach etc.).
Possession Game	No goals/targets in which retention of possession rather than scoring is the primary objective.
Conditioned Game	Restrictions/variations to rules, goals, or area of play, but with both teams scoring in the same way.
Small Sided Game	Two goals, realistic to regulation rules, with both teams scoring in the same way.
Full Sided Game	Two goals, regulation rules and players (i.e. 11 vs. 11 in case of field hockey), with both teams scoring in the same way.
Management/Transition States	
State – Other*	Time when coaches are managing/addressing the team of players to explain up-and-coming practices, transition or move players from one practice state (i.e. technical practice) to a NEW or DIFFERENT practice state (i.e. skills practice), and time when players are taking a breaks from practice (i.e. water breaks).
Match/Game States	
Performance States within Match Competition	<i>Performance (includes 1st Half/Q^{ter}/Game, 2nd Half/Q^{ter}/Game, 3rd Q^{ter}/Game, 4th Q^{ter}/Game)</i> Coaching behaviours that occur during performance (i.e. re-enforcement of good play, verbal or non-verbal scolds etc.).
Non-Performance States within the Match competition	
Timeout	Coaching behaviours that occur during a timeout (i.e. instruction, re-enforcement of good play, verbal or non-verbal praise, verbal or non-verbal scolds etc.).
Half-Time	Coaching behaviours that occur during the half-time break in hockey and basketball (i.e. instruction, re-enforcement of good play, verbal or non-verbal praise, verbal or non-verbal scolds etc.).
End of Quarter/ Game	Coaching behaviours that occur during an end of a quarter in basketball or the conclusion of one of the games in volleyball (i.e. instruction, re-enforcement of good play, verbal or non-verbal praise, verbal or non-verbal scolds etc.).
Non-Performance States outside of the formal Match competition	
Pre-Match	Coaching behaviours that occur before performance commences (i.e. during the pre-match team talk etc.).
End of Match	Coaching behaviours that occur at the conclusion of the whole match/competition (i.e. instruction, feedback of good play, verbal or non-verbal praise, verbal or non-verbal scolds etc.).

Procedure. Three sports were coded by 3 coders [$n = 9$ training sessions ($n = 3$ hockey; $n = 3$ basketball; $n = 3$ volleyball) and $n = 6$ games ($n = 2$ hockey; $n = 2$ basketball; $n = 2$ volleyball)] undergoing observer reliability checks. Reliability was established on a 30-minute section of video from coaching sessions on one of these three sports. Sessions for observer reliability testing were randomly chosen.

Individual sessions were coded by one single coder, for example, coder A. Once this session had been coded and after a one-week time-period (Darst et al., 1989), the same coder completed an intra-observer reliability check on this 30-minute section of video.

This was independent from, and in addition to, an inter-observer reliability check on the same section of video by another coder (i.e. coder B). Figure 2 illustrates the combinations of coders utilised in this ‘triangular’ arrangement. This meant that in the 15 sessions, a given coder would have recorded behaviours for 10 original sections of tape and also undertaken an intra-observer test for these sections of video. The same coder would then have also conducted inter observer tests for ten, 30-minute sections of video across the three sports. Thus, each 30-minute section of video would have generated one intra-reliability score and one inter-observer reliability score.

Table IV. Secondary detail of behaviours: Recipient (Step 3).

Behavioural Classification <u>Recipient</u>	Behavioural Description
Individual	Coach talks or responds non-verbally to a single player either one to one, by pulling a player aside during practice or in a match/competition, uses a player's name, etc. in BOTH PRACTICE and in MATCH/COMPETITION contexts
Group	IN PRACTICE GROUP refers to the coach talking or responding non-verbally to more than one player, up to half of the team of players. IN MATCH/COMPETITION GROUP refers to more than one player on the court/field of play, and players who are substitutes who are NOT on the field of play, irrespective of whether these substitutes are greater in number than those on the court/field on play.
Team	IN PRACTICE TEAM refers to the coach talking or responding non-verbally to more than one half of the team of players. IN MATCH/COMPETITION TEAM refers to those players on the court/field of play.
Other	Coach talks or responds non-verbally to an assistant coach, referee, parent, spectator, etc.

Table V. Secondary detail of behaviours: Timing (Step 4).

Behavioural Classification <u>Timing</u>	Behavioural Description
Pre	Information given BEFORE a performance episode
Concurrent	Information given DURING a performance episode.
Post	Information given AFTER a performance episode.

Table VI. Secondary behaviours: Content (Step 5).

Behavioural Classification <u>Content</u>	Behavioural Description
Technical:	e.g. related to individual techniques
Technically-based	such as passing, shooting, dribbling, etc.
Tactical: Cognitive strategies that involve player decision-making	e.g. patterns of play, formations (e.g. defensive shape), counter attacking, pressing, player movement and connections etc.
Other	Not fitting any other behaviour category.

Once the 30-minute section of video had been coded for one individual session and the intra- and inter-observer reliability completed, levels of agreement were calculated, this was repeated after each re-coding session. If the calculated level of agreement did not match the required level, then coders undertook more training. This consisted of clarifying reasons for disagreement and reviewing the behavioural definitions to encourage procedural reliability. Common patterns of error were noted. These were coders missing behaviours when they occurred simultaneously to each other (i.e. corrective feedback and positive/negative modelling), when behaviours were sandwiched together (i.e. general feedback

Table VII. Secondary behaviours: Questioning and silence (Step 6).

Behavioural Classification <u>Questioning</u>	Behavioural Description
Convergent	Limited number of correct answers/options – closed responses (i.e. often yes or no answer). e.g. 'Which is the best passing option from here, forward or back?', 'Did you get your injury treated?', 'are you ready to sub on the field of play?', 'are you going to call play one?', 'why did you do that?', 'do you think that option was good one?', 'was that the way that I showed you to do that skill?', 'was that the correct choice?', 'which team is defending?', 'do you understand what I want you to do?', 'can you show me what you just did?'
Divergent	Multiple responses/options – open to various responses. e.g.: 'what options do you have available in that position?', 'where were you looking to see whether the spike was coming straight or cross court?', 'when do you think it might be a good idea to step up and press the ball?', 'how might you help team-mates who are out of position defensively?', 'what things might you do to help the player in possession of the ball?', 'once we have committed all the defenders into that channel, where do we go?'
Silence	
Silence on-task	Coach monitors practice without reacting verbally or non-verbally.
Silence off-task	Coach is not visibly engaged in the practice.

which was followed by corrective feedback), and when the same behaviour occurred twice in a row directed to different recipients as discrete behaviours (i.e. the coach instructed the team, and then followed this up with instruction to an individual).

After follow-up training, the coding intra- and inter-reliability tests were repeated for all 15 sessions, and levels of agreement were re-calculated. Agreements had to be made between all three coders for

Sport	Session	Coder		
		A	B	C
Hockey	Session 1	√		√
	Session 2	√	√	
	Session 3		√	√
	Session 4	√		√
	Session 5	√	√	
Basketball	Session 1	√	√	
	Session 2		√	√
	Session 3	√		√
	Session 4	√	√	
	Session 5		√	√
Volleyball	Session 1		√	√
	Session 2	√		√
	Session 3	√	√	
	Session 4		√	√
	Session 5	√		√
Total		10	10	10

Figure 2. Figure outlining the coding 'triage' employed during Stage 5 of the development and validation of the CAIS.

each of the 15 sessions and not simply two coders. This meant there was more difficulty in reaching levels of agreement. In addition, while levels of intra-observer agreement may have been reached for a 30-minute section of video, inter-observer may have not, so the process mentioned above was repeated until all 15 sessions met appropriate levels of agreement.

The secondary match/game and/or training 'state' behaviours were completed on the 15 coaching sessions selected for intra- and inter-observer reliability checks. The durations of these sessions ranged from 54–87 minutes for matches/games ($n=6$) and 71–120 minutes in length for training/practice

sessions ($n=9$). Completing the event duration recording analysis on the full length of the 15 coaching sessions for the secondary match/game and/or training 'state' behaviours and NOT just the 30-minute sections of video ensured that the full range of match/game and training/practice behaviours were coded for intra- and inter-observer reliability. Due to differences in the match/game and/or training 'state' behaviours that required coding (see Table I), analysis of intra- and inter-observer agreement levels was conducted separately for match/games durations ($n=6$) and training/practice session durations ($n=9$).

The coding of the remaining secondary detail of behaviours associated with steps 3–6 of the CAIS (see Figure 1) was completed alongside coding of primary behaviours and, therefore, followed the same procedures in reaching appropriate levels of agreement. The only difference was that the process was made easier as the amount of secondary detail of behaviours in each category (i.e. recipient, timing, content, questioning and silence) was far less than the 23 primary behaviours. What is more, the fact that secondary detail is always linked to a primary behaviour and could not occur when not linked to the primary behaviour meant issues such as ‘missing’ the behaviour were far less likely to occur and levels of agreement were much higher after the first iteration. So, when a primary behaviour was missed by a coder, this meant that the secondary detail of behaviours could not be attached. This therefore resulted in some differing levels of agreement for some of the secondary detail of behaviours. However, some definitions of secondary detail of behaviours did need clarifying, especially in respect to the distinction between when the coach addressed the team and group in match play ‘performance’ settings. Clarification was given concerning competition definitions. For example, the ‘team’ was linked to the players on the field of play while the substitutes on the side waiting to enter the game represented the ‘group’.

Results. Once all the data were coded both the primary and secondary detail of behaviours were reviewed and the intra- and inter-observer reliability coefficients using the agreements/(agreements + disagreements) method (van der Mars, 1989) were calculated. There is no general standard on agreement levels for behavioural research, with 80–85% deemed sufficiently high, but agreement level is dependent on the complexity of the instrument (Johnston & Pennypacker, 2009; van der Mars, 1989). Given the complexity of the CAIS, reliability levels were set at the acceptable standard of 80% (Gilbert et al., 1999; van der Mars 1989). In an additional step, completing four two-way analysis of variance tests, one for primary behaviours, one for the match/game ‘state’ durations, one for the training/practice durations, and one for the secondary behaviours further assessed observer agreement in this stage of the development and validation. The aim of these tests was to not only assess for relationships between the observers but also to test for individual variability in coding which would be revealed by any statistically significant differences between observers over the course of the different sessions they coded (Nevill, 1996; Nevill, Lane, Kilgour, Bowes, & Whyte, 2001).

For the primary behaviours average intra-observer agreement reached the critical level of 0.80 (0.82)

across the sessions ($n=15$). Coefficients ranged from 0.75–0.88. Average inter-observer agreement for the primary behaviours also reached the critical level of 0.80 (0.81), with inter-observer reliability coefficients ranging from 0.76–0.86.

Some of the coefficients were below the 0.80 average. In terms of the primary behaviours, only two of the 15 sessions coded for intra-observer reliability fell below the critical level of 0.80 (one at 0.75 and one at 0.78) and only two of the sessions for inter-observer reliability (one at 0.76 and one at 0.78) fell below the critical level. The errors made derived from observer drift and/or misclassification (on both primary and secondary behaviours). These were issues that were dealt with corrective coder training. Given the depth of the instrument, issues of procedural reliability occurred. Most issues originated from the primary behaviours. In these instances, one coder may have assigned two codes to a primary intervention behaviour where another other coder only assigned one to this same incident and vice versa. For example, a coach may have given two ‘instructions’ (i.e. ‘get behind the ball’ to an individual and followed this by shouting ‘press’ to the group/team). Some coders had coded this correctly as two separate behaviours, while others had coded it as a single behaviour. Similar issues occurred when the coaches utilised multiple intervention behaviours concurrently. For example, the coach may have given corrective feedback to an individual but at the same time used a negative model. Some coders missed one of the two intervention behaviours simply treating this as one behaviour, instead of separating out each incident.

In addition to the primary behaviour observer agreement correlation statistics, results from the two-way analysis of variance test showed that there were no statistically significant differences for observer ($F=1.653$, $df\ 2$, $P=0.203$), session ($F=0.888$, $df\ 3.989^1$, $P=0.380$), or the interaction of observer and session ($F=0.759$, $df\ 28$, $P=0.811$) across the 15 sessions coded. These results suggest that differences in data coding were not attributable to individual variability (Williams, Ward, Smeeton, & Allen, 2004).

For the secondary match/game ‘state’ behaviours average intra-observer agreement reached the critical level of 0.80 (0.98) across the sessions ($n=6$). Coefficients ranged from 0.96–1.00 for event recording durations. Average inter-observer agreement for the secondary detail of training/practice ‘state’ behaviours also reached the critical level of 0.80 (0.92) across the sessions ($n=9$), with inter-observer reliability coefficients ranging from 0.82–0.99 for event recording durations. As with the primary behaviour coding, the two-way analysis of variance test conducted on the match/game ‘state’ duration behaviours were not statistically significant for observer ($F=0.000$, $df\ 1.079^1$, $P=1.000$), session

($F = 0.855$, $df 2.258^1$, $P = 0.454$) or the interaction of observer and session ($F = 0.000$, $df 10$, $P = 1.000$). In addition, the two-way analysis of variance test conducted on the training/practice 'state' duration behaviours were not statistically significant for observer ($F = 0.000$, $df 1.247^1$, $P = 1.000$), session ($F = 0.192$, $df 2.543^1$, $P = 0.873$) or the interaction of observer and session ($F = 0.000$, $df 16$, $P = 1.000$). These results again suggest that differences in data coding were not attributable to individual variability.

For the remaining secondary detail of behaviours (i.e. recipient, content and timing) average intra-observer agreement reached the critical level of 0.80 (0.89) across the sessions ($n = 15$). Coefficients ranged from 0.79–0.99. Average inter-observer agreement for the remaining secondary detail of behaviours (i.e. recipient, content and timing) also reached the critical level of 0.80 (0.90), with inter-observer reliability coefficients ranging from 0.77–0.97. In only one of the 15 sessions was intra- and inter-observer reliability below the critical level (0.79 and 0.77, respectively). As with the previous analyses, the two-way analysis of variance test conducted on the secondary behaviours did not reveal any statistically significant results for observer ($F = 0.019$, $df 1.402^1$, $P = 0.947$), session ($F = 1.451$, $df 5.238^1$, $P = 0.204$), or the interaction of observer and session ($F = 0.057$, $df 28$, $P = 1.000$). These results again suggest that differences in data coding were not attributable to individual variability.

Discussion

We developed and validated a computerised, multi-dimensional, hierarchical analysis system able to identify specific behaviours that occur within complex coaching environments. Specifically, we described the evolution of the Coach Analysis and Intervention System (CAIS) and the process for establishing the validity and reliability of it. Following the validation procedure the completed instrument developed for use in practice sessions and competition had 23 primary behaviours related to physical behaviour, feedback/reinforcement, instruction, verbal/non-verbal, questioning and management. The instrument also considered secondary detail of behaviour related to performance states, content, recipient, and timing. The resulting instrument is a multidimensional and multilevel mechanism able to provide detailed and contextualised data about specific coaching behaviours occurring in complex and nuanced coaching interventions and environments.

It is hoped that the CAIS will contribute to the field through its ability to clearly differentiate behavioural classifications and enable consideration of the dimensions of that behaviour. Not only

through identifying the specifics of the behaviour (i.e., the nature and content), but its temporal patterning (i.e., when it occurs in relation to performance, and in what 'state'), and its direction (i.e., the recipient). These are important considerations as differences in coaching behaviour in coaching interventions will conceivably have a differential effect upon performance (Brewer & Jones, 2002), as will the proportions of time engaging in different practice states, and interaction of both (Ford et al., 2010). In turn, analysing behaviour with some context (content, timing, stage of practice/performance, recipient) the instrument allows an examination of the microstructure of coach behaviour and discernable patterns of 'coaching' to be identified. These patterns may be linked to established understandings of skill learning and the activities and behaviours that develop expertise i.e. the timing and content of feedback linked with practice that is random and variable (Ford et al., 2010; Williams & Hodges, 2005). Moreover, identifying patterns of behaviour allow informed judgements about coaches' behaviour in specific contexts. In addition, a more encompassing systematic observation system can assist in an expert performance approach to coaching where detailed categorisation helps to capture the performance of coaches in as much detail as possible, and the application procedures for establishing valid instruments strengthen the outcomes from this line of enquiry (Ford, Coughlan, & Williams, 2009).

The trade-off for detail of analysis is complexity, with numerous dimensions and behavioural categories. However, given the goal of observational recording of coaching environments is to produce an accurate reflection of the events that occurred (Brewer & Jones, 2002; Darst et al., 1989), this instrument adds to a more complete and accurate picture of coaching interventions. Moreover, through the methodology described, high levels of familiarisation with the behaviour categories along with observer agreement levels demonstrated a reducing effect on instrument complexity. Nonetheless, given the complex nature of the instrument, there remains a need for rigorous training for potential users in the behavioural definitions and use of the tool, particularly with those unfamiliar with systematic observation (Brewer & Jones, 2002). In addition, while this research reports the process to develop and validate the CAIS in its entirety, a utility of the instrument is its ability to allow researchers to focus on specific aspects of coach behaviour, using one or two elements of primary behaviour (Table II) with associated detail of secondary behaviour (Tables III–VII) according to need. In this way the CAIS can be used as a series of 'sub-instruments' to focus on specific aspects of research, coaching practice or coach education and

development. Given the continued developments in computer technology, this approach lends itself to live coding with a hand-held or tablet device, giving the coach instant feedback, and allowing the live coding to be linked back and synched with the recorded video footage for further and fuller analysis. The authors are currently developing this approach.

Moving the literature and its application forward requires researchers more able to capture the complexity that is inherently linked to the context in which coaching occurs. It is hoped that the CAIS can assist with this. It is of course recognised that coaching behaviours *per se* do not stand alone as predictors of effective coaching (Dodge & Hastie, 1993) nor do they “embrace the entirety of the coaching process” (Lyle, 1999, p.14). Combined methodologies are increasingly employed (e.g., Potrac, Jones, & Armour, 2002; Potrac et al., 2007), with systematic observation combined with interviewing revealing much about coaches cognitive strategies and the context of behaviour. More detailed longitudinal research with tools sensitive to the nuances involved in coaching might be better able to explain the events, people, and mediating factors that comprise the multifarious coaching process. From the resulting knowledge, coach education can utilise meaningful interventions.

The hierarchical nature of the CAIS illustrates something of the steps involved in the recording of coaches' behaviours beyond the 'what'. Observing and analysing in this way can demonstrate that feedback or instruction is 'technical' during training but 'tactical' during match play or that although the coach is using questioning, it is all directed to the team and is convergent in nature (i.e. closed questions are being employed by the coach). This allows for richer descriptions of coach behaviour so that interventions focused on improving interaction can be targeted to specific areas of practice. In turn, this develops coaches' self-awareness and enables questions about congruence between actual behaviour and stated intentions to be considered (Cushion, 2010), facilitating for coach education an alignment of practice with practice statements, and matching knowledge and action. This analysis coupled with scaffolded reflection can enable coaches to identify coach education needs, and assists in taking significant steps towards guiding coaches to effective pedagogical practices taking account of their daily coaching realities.

Acknowledgements

The authors would like to thank the reviewers and editors for their helpful comments in developing this paper. We would also like to thank the coders for their assistance in this research.

References

- Abraham, A., Collins, D., Smethurst, C., & Collins, C. (1997). The use of coach behavioural assessment and associated problems. *Journal of Sports Sciences*, 15, 70.
- Baumgartner, T.A., Jackson, A.S., Mahar, M.T., & Rowe, D.A. (2007). *Measurement for evaluation in physical education and exercise science* (8th ed.). New York, NY: McGraw-Hill.
- Bloom, G.A., Crumpton, R., & Anderson, J.E. (1999). A systematic observation study of the teaching behaviours of an expert basketball coach. *The Sport Psychologist*, 13, 157–170.
- Bloom, G.A., Durand-Bush, N., & Salmela, J.H. (1997). Pre- and post competition routines of expert coaches of team sports. *The Sport Psychologist*, 11, 127–141.
- Brewer, C.J., & Jones, R.L. (2002). A five-stage process for establishing contextually valid systematic observation instruments: The case of rugby union. *The Sport Psychologist*, 16, 139–161.
- Cheffers, J. (1977). Observing teaching systematically. *Quest*, 28, 17–28.
- Côté, J., Salmela, J.H., Trudel, P., Baria, A., & Russell, S. (1995). The coaching model: A grounded assessment of expert gymnastic coaches' knowledge. *Journal of Sport & Exercise Psychology*, 17, 1–17.
- Cushion, C.J. (2007). Modelling the complexities of the coaching process. *International Journal of Sports Science and Coaching*, 2, 395–401.
- Cushion, C.J. (2010). Coach Behaviour. In J. Lyle & C.J. Cushion (Eds.), *Sports coaching professionalization and practice* (pp. 243–253). London: Elsevier.
- Cushion, C.J., & Jones, R.L. (2001). A systematic observation of professional top-level youth soccer coaches. *Journal of Sport Behaviour*, 24, 1–23.
- Darst, P.W., Zakrajsek, D.B., & Mancini, V.H. (1989). Preface. In P.W. Darst, D.B. Zakrajsek, & V.H. Mancini (Eds.), *Analyzing physical education and sport instruction* (2nd ed.) (pp. ix–xi). Champaign, IL: Human Kinetics.
- DeMarco, G.M.P., Mancini, V.H., & Wuest, D.A. (1996). Reflections on change: A qualitative and quantitative analysis of a baseball coach's behaviour. *Journal of Sport Behaviour*, 20, 135–163.
- Dodge, B., & Hastie, P. (1993). Coach effectiveness. *Sport Science Review*, 2, 14–29.
- Ford, P.R., Coughlan, E., & Williams, M. (2009). The expert performance approach as a framework for understanding and enhancing coaching performance expertise and learning. *International Journal of Sport Science and Coaching*, 4, 451–463.
- Ford, P.R., Yates, I., & Williams, M.A. (2010). An analysis of practice activities and instructional behaviours used by youth soccer coaches during practice: Exploring the link between science and application. *Journal of Sports Sciences*, 28, 483–495.
- Franks, I., Johnson, R., & Sinclair, G. (1988). The development of computerized coaching analysis system for recording behaviour in sporting environments. *Journal of Teaching in Physical Education*, 8, 23–32.
- Gilbert, W.D., & Trudel, P. (1999). An evaluation strategy for coach education programs. *Journal of Sport Behavior*, 22, 234–250.
- Gilbert, W., & Trudel, P. (2004). Analysis of coaching science research published from 1970–2001. *Research Quarterly for Exercise and Sport*, 75, 388–399.
- Gilbert, W., Trudel, P., Gaumond, S., & Larocque, L. (2004). Development and application of an instrument to analyse coaches' pedagogical content interventions. *Sociology of Sport On Line*, 2(2). Retrieved from <http://www.physed.otago.ac.nz/sosol/v2i2/v2i2.htm>.
- Hawkins, A., Sharpe, T.L., & Ray, R. (1994). Toward instructional process measurability: An interbehavioural field systems perspective. In R. Gardener, D.M. Saintano, J.O. Cooper, T.E. Heron, W.L. Heward, J. Eshleman, & T.A. Grossi (Eds.), *Behaviour analysis in education: Focus on measurably superior instruction* (pp. 241–255). Pacific Grove, CA: Brooks/Cole.

- Hodges, N. J., & Franks, I.M. (2002). Modelling coaching practice: The role of instruction and demonstration. *Journal of Sports Sciences*, 20, 793–811.
- Horn, T.S. (2002). Coaching effectiveness in the sport domain. In T.S. Horn (Ed.), *Advances in sport psychology* (pp. 309–354). Champaign, IL: Human Kinetics.
- Johnston, J.M., & Pennypacker, H.S. (2009). *Strategies and tactics of behavioral research* (3rd ed.). New York, NY: Routledge.
- Kahan, D. (1999). Coaching behaviour: A review of the systematic observation research literature. *Applied Research in Coaching and Athletics Annual*, 14, 17–58.
- Lacy, A.C., & Darst, P.W. (1984). Evolution of a systematic observation instrument: The ASU Observation instrument. *Journal of Teaching in Physical Education*, 3, 59–66.
- Lacy, A.C., & Goldston, P.D. (1990). Behaviour analysis of male and female coaches in high school girls' basketball. *Journal of Sport Behaviour*, 13, 29–49.
- Langsdorf, E.V. (1979). *Systematic observation of football coaching in a major university environment*. (Unpublished master's thesis). Arizona State University, USA.
- Lombardo, B.J., Faraone, N.F., & Pothier, D. (1983). *The coach in action: A descriptive analysis*. Paper presented at the convention of the Eastern district association of the American Alliance for Health, Physical Education, Recreation and Dance, Providence, RI.
- Lyle, J. (1999). Coaching philosophy and coaching behaviour. In N. Cross & J. Lyle (Eds.), *The coaching process: Principles and practice for sport* (pp. 25–46). Oxford: Butterworth-Heinemann.
- Lyle, J., & Cushion, C.J. (2010). Narrowing the field: Some key questions about sports coaching. In J. Lyle & C.J. Cushion (Eds.), *Sports coaching professionalization and practice* (pp. 243–253). London: Elsevier.
- Mesquita, I., Rosado, A., Januario, N., & Barroja, E. (2008). Athlete's retention of a coach's instruction before a judo competition. *Journal of Sport Science and Medicine*, 7, 402–407.
- Metzler, M. (1989). The ALT-PE microcomputer data collection system. In P.W. Darst, D.B. Zakrajsek, & V.H. Mancini (Eds.), *Analyzing physical education and sport instruction* (2nd ed.) (pp. 225–232). Champaign, IL: Human Kinetics.
- More, K. (2008). Notational analysis of coaching behaviour. In M. Hughes & I. Franks (Eds.), *The Essentials of Performance Analysis* (pp. 264–276). London: Routledge.
- Nash, C., & Collins, D. (2006). Tacit knowledge in expert coaching: Science or art? *Quest*, 58, 465–477.
- Nevill, A.M. (1996). Validity and measurement agreement in sports performance [Editorial]. *Journal of Sports Sciences*, 14, 199.
- Nevill, A.M., Lane, A.M., Kilgour, L.J., Bowes, N., & Whyte, G.P. (2001). Stability of psychometric questionnaires. *Journal of Sports Sciences*, 19(4), 273–278.
- Pereira, F., Mesquita, I., & Graça, A. (2009). Accountability systems and instructional approaches in youth volleyball training. *Journal of Sport Science and Medicine*, 8, 366–373.
- Potrac, P., Jones, R.L., & Armour, K. (2002). 'It's all about getting respect': The coaching behaviours of an expert English soccer coach. *Sport Education and Society*, 7, 183–202.
- Potrac, P., Jones, R.L., & Cushion, C.J. (2007). Understanding power and the coach's role in professional English soccer: A preliminary investigation of coach behaviour. *Soccer and Society*, 8, 33–49.
- Quartermann, J. (1980). An observational system for observing the verbal and nonverbal behaviors emitted by physical educators and coaches. *The Physical Educator*, 37, 15–20.
- Rushall, B.S. (1977). Two observational schedules for sporting and physical education environments. *Canadian Journal of Applied Sports Sciences*, 2, 15–21.
- Saffrit, M.J. (1986). *Introduction to measurement in physical education and exercise science*. St-Louis, MO: Times Mirror/Mosby.
- Schempp, P., McCullick, B., St. Pierre, P., Woorons, S., You, J.A., Clark, B. (2004). Expert golf instructors' student-teacher interaction patterns. *Research Quarterly for Exercise and Sport*, 75, 60–70.
- Sherman, M.A., & Hassan, J.S. (1984). Behavioural studies of youth sport coaches. In M. Pieron & G. Graham (Eds.), *The 1984 Olympic scientific congress proceedings: Sport pedagogy* (pp. 103–108). Champaign, IL: Human Kinetics.
- Sinclair, G.D. (1989). Feedback analysis profile (FAP). In P.W. Darst, D.B. Zakrajsek, & V.H. Mancini (Eds.), *Analyzing physical education and sport instruction* (2nd ed.) (pp. 361–368). Champaign, IL: Human Kinetics.
- Smith, M., & Cushion, C.J. (2006). An investigation of the in-game behaviours of professional, top-level youth soccer coaches. *Journal of Sports Sciences*, 24(4), 355–366.
- Smith, R.E., Smoll, F.L., & Hunt, E.B. (1977). A system for the behavioural assessment of athletic coaches. *Research Quarterly*, 48, 401–407.
- Stratton, G., Reilly, T., Williams, A.M., & Richardson, D. (2004). *Youth soccer: From science to performance*. London: Routledge.
- Strean, W. (1995). Youth sport contexts: Coaches' perceptions and implications for intervention. *Journal of Applied Sport Psychology*, 7, 23–37.
- Tannehill, D., & Burton, D. (1989). Coaching behaviors observational recording system. In P.W. Darst, D.B. Zakrajsek, & V.H. Mancini (Eds.), *Analyzing physical education and sport instruction* (2nd ed.) (pp. 379–389). Champaign, IL: Human Kinetics.
- Tharp, R.G., & Gallimore, R. (1976). What a coach can teach a teacher. *Psychology Today*, 9, 75–78.
- Trudel, P., Côté, J., & Bernard, D. (1996). Systematic observation of youth ice hockey coaches during games. *Journal of Sport Behaviour*, 19, 50–65.
- van der Mars, H. (1989). Observer reliability: Issues and procedures. In P.W. Darst, D.B. Zakrajsek, & V.H. Mancini (Eds.), *Analyzing physical education and sport instruction* (2nd ed.) (pp. 53–80). Champaign, IL: Human Kinetics.
- Vogt, W.P. (1999). *Dictionary of statistics and methodology* (2nd ed.). Thousand Oaks, CA: Sage.
- Williams, A.M., Ward, P., Smeeton, N.J., & Allen, D. (2004). Developing anticipation skills in tennis using on-court instruction: Perception versus perception and action. *Journal of Applied Sport Psychology*, 16, 350–360.
- Williams, A.M., & Hodges, N.J. (2005). Practice, instruction and skill acquisition: Challenging tradition. *Journal of Sports Sciences*, 23, 637–650.