

This is the accepted manuscript of Cope, E., Partington, M. & Harvey, S. (2016) A review of the use of a systematic observation method in coaching research between 1997-2016, *Journal of Sports Sciences*, published on 18 Oct 2017 at <https://doi.org/10.1080/02640414.2016.1252463>

^a*Department of Sport, Health and Exercise Science, University of Hull, Cottingham Road, Hull, East Yorkshire, UK, HU6 7RX, Ed.Cope@hull.ac.uk;* ^b *Department of Sport and Physical Activity, Edge Hill University, Ormskirk, UK, parting@edgehill.ac.uk;* ^c *College of Physical Activity and Sport Sciences, West Virginia University, 375 Birch Street, Morgantown, West Virginia, USA, Stephen. Harvey@mail.wvu.edu*

Abstract

A systematic observation method has been one of the most popularly employed methods in coaching research. Kahan's review of this method conducted between 1975-1997 highlighted the key trends in this research, and offered methodological guidance for researchers wishing to use this method in their research. The purpose of this review was to provide an update of the use of a systematic observation method in coaching research and assess the extent to which the calls made by Kahan have been addressed. While in some respect this field of study has progressed (i.e. the introduction of qualitative methods), researchers adopting this method have failed to attend to many of the issues Kahan raised. For this method to continue to make a positive contribution toward the coaching research literature, researchers need to more critically reflect on how and why they are employing this method. At present, some of the decisions made by researchers who have conducted work in this area are not justified with a rationale. It is our intention that this review will serve as guidance for researchers and practitioners, and editors and reviewers of journals when attempting to assess the quality of this type of work.

Keywords

Systematic observation; coaching; coaching behaviour; methodology; coaching practice

Introduction

A previous review of studies using systematic observation methods in coaching (Kahan, 1999) included 56 studies that had used this method to observe coaching behavior during the period of 1975-1997. Moreover, a review of coaching science research from 1970-2001 (Gilbert & Trudel, 2004) revealed that the study of coaching behavior was the main area under investigation, with 13.1% of all studies included in this review using a systematic observation method. Given these figures, it is clear that the coaching research community sees systematic observation as a valuable tool in developing a greater understanding of what coaches do in practice and competition.

In his review, Kahan (1999) raised some concerns with research that had employed systematic observation: 1) that studies were mostly conducted from a positivistic perspective and so rarely considered the contextual factors that impacted coaches' behavior; 2) studies had been conducted in a small number of sports (i.e. basketball, football and soccer), and mostly in a youth sport context; 3) few studies observed coaches' behavior within training and competition; 4) sample sizes were small and often not randomly sampled; and, 5) conclusions of coaches' behavior were made based on a limited number of observations that only produced a 'snapshot' of those coaches practices. On a positive note, Kahan (1999) suggested systematic observation in coaching has revealed a lot about what coaches do, although judgments of the appropriateness of this behavior were unable to be made due to a limited knowledge of factors related to athletes (i.e., their learning needs and motivations for participating), and the context in which they participated (i.e., what coaches were attempting to achieve related to the context).

In the period since Kahan's review, the use of systematic observation has remained popular amongst coaching researchers, and has continued to evolve as the field of coaching has become more established. However, the extent to which this evidence has contributed toward the development of coaching practice, especially within the confines of the specific contexts in which 'coaching' takes place (Lyle, 2002), is unclear. Therefore, the purpose of this paper is to present an updated review of research into the use of a systematic observation method to record coaching behavior, and consider how this line of research has moved forward since Kahan's review. First, we overview how we identified studies to be included in the review.

Method

Identification of studies

Coaching studies using a systematic observation method were searched using a three-phase approach (Harvey & Jarrett, 2006). Phase one involved searching the EBSCO HOST database. Specific databases searched were Academic Search Complete, Educational Research Complete, ERIC, PsycArticles, PsycBooks, PsycInfo, and SPORTdiscus with FullText. Original search terms followed those of Kahan (1999), which were *systematic observation AND coaching AND behaviour*. Closely related terms and those used in studies that were known to have used a systematic method, such as *coach* and *athlete* and *learning* were also included in searches to ensure all relevant articles that met the inclusion criteria were identified. Database searches stopped once a saturation point had been reached, which was when no new articles were found.

Phase two expanded the search beyond the databases to involve other studies that met the inclusion criterion: post 1997, empirical, peer-reviewed study, written in English, the participants of the study were coaches, and a category-based, systematic observation instrument to observe coaching behaviour directed toward players. This extended search was achieved by reading the reference lists of articles identified in phase one, as well as emailing researchers who were known to conduct coaching research using a systematic observation method. Finally, colleagues directed the authors to any other studies that had not been identified through any other means. Any studies that did not meet the inclusion criteria were removed. These were studies that were of a theoretical nature, or with the purpose of validating a systematic observation instrument, and focused on teachers rather than coaches.

To ensure reliability, a three-step process, as outlined and implemented by Gilbert and Trudel (2004) and LaVoi and Dutove (2012) was followed. First, all members of the research team agreed to the criteria for article inclusion. Once it was agreed that the study should be included, it was allocated to a member of the research team to read and code. Second, the first and second authors drew upon the experience and expertise of the third author, who had been trained in and published similar work, for guidance on coding articles included for review. Finally, the first and second author coded 25% of the articles (n=6) independently from a random sample of articles. Inter-coder reliability was 96%, with the one disagreement discussed until consensus was obtained.

Summary of Studies from 1997 to 2016

Twenty-six studies on the use of a systematic observation method in coaching were identified in the current review. To document the information from each study, a coding system was designed. Initial categories of this coding system were informed by: 1) previous reviews of coaching behavior to allow comparisons to be made (Kahan, 1999; Trudel, Côté, J, & Bernard, 1996) and, 2) the authors' experiences of conducting similar reviews in coaching. For each study the following categories were coded: a) sports; b) countries, c) coaching context; d) systematic observation instrument; e) additional methods; f) number of total observations per coach; g) observation frequencies across studies; h) method of recording; i) reliability procedure. Coding information for each of these categories resulted in these being combined into four broader themes: 1) instrument development and technology, 2) coder training, reliability, and procedural issues, 3) research questions and paradigm shift, 4) research context.

Results

Instrument development and technology

Researchers employed a range of systematic observation instruments. The most common was the Arizona State University Observation Instrument (ASUOI), which was used in nine studies. However, four of these studies used a modified or adapted version, rather than Lacy and Darst's (1984) original version. A similar story existed with the Coach Analysis Intervention System (CAIS) as six of the eight studies used a modified or adapted version rather than Cushion, Harvey, Muir and Nelson (2012) validated version. The Coach Behavior Recording Form (CBRF) was used in three studies but again on each of these occasions a modified or adapted version was employed. The System for Observing the Teaching of Games in Physical Education (SOTG-PE), the Rugby Coach Activities and Behavior Instrument (RCABI), the Coach Behavior Assessment System (CBAS), the Cheffers Adaptation of Flanders Interaction Analysis System (CAFIAS), and the Coach Athlete Interaction Coding System (CAICS) were used in only one study each. In each of these cases, the full versions of the instruments were used, apart from the SOTG-PE which was modified in Vinson, Brady, Moreland, and Judge (2016) study. Four studies took components from more than one instrument to create a new, hybrid instrument.

Table 1. Systematic observation instruments used to study coaching behaviour

Instruments	Frequency (number of which were modified)
ASUOI	9 (4)
CAIS	8 (6)
CBRF	3 (3)
SOTG-PE	1 (1)
CBAS	1 (1)
RCABI	1
CAFIAS	1
CICS	1
Hybrid system	4
Total	29

Note: The total equals 29 because some studies employed more than one systematic observation method.

Coder training, reliability, and procedural issues

Coder training, and intra and inter-observer reliability scores if specified, were recorded for each study. Seventeen studies indicated that coders had been trained in using the systematic observation instrument employed with one study (Becker & Wrisberg, 2008) stating that consensus training had taken place. Seven studies failed to report if any coder training had taken place. Furthermore, Seventeen studies provided intra-observer reliability scores, while seven did not, with twenty studies providing inter-observer reliability scores where as five did not.

Table 2. Coder training, inter and intra reliability

Procedure	Number of studies (number of additional studies not reported)
Coder training	18 (7)
Consensus building technique	1
Inter reliability	20 (5)
Intra reliability	18 (7)

Note: The total equals 25 because one study employed a consensus building technique

There were differences in the number of coaches observed for each study depending on its purpose and nature. However, not all coaches were observed within the same study the same number of times. For example, in Harvey, Cushion, Cope and Muir (2013) study, the three coaches were observed a different number of times each. In these instances we grouped coaches together and reported the mean number of observations per study. In six studies, coaches were observed only once, in seven studies coaches were observed for an average of between two to four times, in six studies coaches were observed for an average of between five to seven times, in three studies coaches were observed for an average of between eight to ten times, and in only one study a coach was observed on more than ten occasions. Two studies did not report the number of times coaches were observed, while one study was incalculable.

Table 3. The mean number of coaching observations per coach

Observation frequency	Number of studies
1	6
2-4	7
5-7	6
8-10	3
10>	1
Incalculable	1
Not reported	2
Total	26

Note: Incalculable: studies in which a set datum value across coaches was reported for observation frequency.

Reviewing the number of minutes' coaches were systematically observed was difficult with studies tending to report different descriptions, for example total number of hours, average number of hours and some studies collating the coaches' hours as a group or as individuals. For example, Pereira, Mesquita and Graça, (2009) reported the average session length for all 28 coaches as 87 minutes and the total minutes for all coaches as 2430. Vinson et al. (2016) reported that each case study was systematically observed for approximately 4 hours, up to two hours with two instruments. Because of these indifferences and the challenges in presenting data in a consistent format, we did not report the length in time of observations.

Finally, the method of recording behavior was coded. In ten studies an interval recording method was used, in seven studies time sampled event was used, in five studies event was employed, in three time sampled was used, with five studies failing to report this information.

Table 4. The mean number of coaching observations per coach

Method of recording	Number of studies
Interval	10
Time-sampled event	5
Event	6
Sampled	3
Not reported	5
Total	29

Note: This number equals 29 because some studies used two methods of recording

In addition, nineteen studies reported that they had videoed sessions to allow for post-observation coding, while seven studies had coded behavior live.

Table 5. Method of systematic observation recording

Method of recording	Number of studies
Video	19
Live	7
Total	26

Research questions and paradigm shift

Accompanying a systematic observation instrument, a number of studies used an additional method(s) in an attempt to reveal a further aspect of the coach's practice. Fourteen studies employed some form of interview to predominantly find out the underpinning reasons why coaches used certain behaviors, four studies used time use analysis to find out how long coaches engaged athletes in different practice activities (i.e. technical, phase of play, small-sided game), two studies used observational field notes to uncover descriptive information related to the coach-athlete relationship and coaching context (i.e. how coaches' communicated and how athletes seemed to receive this information), and the modified expectancy rating scale was employed once in order to measure coach expectations of athletes. Eight of the twenty-six studies did not use any additional methods.

Table 6. Additional methods implemented

Method	Frequency
Interview	14
Qualitative observation	2
Focus group	2
Modified expectancy rating scale	1
None	8
Total	27

Note: The total equals 27 because one study employed more than one additional method

Research Context

With respect to the type of sport studied, twenty-nine were team based, which included twelve in football, five in volleyball and basketball, two in handball, and one in rugby union, synchronized swimming, wheelchair basketball, field hockey, and American football. Only one study investigated the behaviors of a coach working in an individual sport, which was golf.

Table 7. Sports reported across studies

Sport	Frequency
Soccer	12
Volleyball	5
Basketball	5
Handball	2

Golf	1
Rugby Union	1
Synchronised Swimming	1
Wheelchair Basketball	1
Field Hockey	1
American Football	1
Total	30

Note: The total equals 30 because some studies used a systematic observation method to investigate coaches' behaviour in more than one sport

There has been an increase in the range of geographical location of systematic observation research in coaching. Results from this review indicate that it is now the UK where most of this research is being completed, with thirteen studies conducted during the review period. Along with this, three studies were undertaken in Portugal, the USA, and Spain, two in Canada, and one in Australia and Greece.

Table 8. Country reported across studies

Country	Number of studies
United Kingdom	13
Portugal	3
United States of America	3
Spain	3
Canada	2
Australia	1
Greece	1
Total	26

Based on Trudel and Gilbert's (2006) conceptualization of coaching contexts, it was identified that six studies had been conducted in a recreational context, twelve in a developmental context, and twelve in an elite context*.

Table 9. Coaching domain reported across studies

Domain	Frequency
Participation	6
Development	12
Performance	12
Total	30

Note: This total equals 30 because some studies used a systematic observation method across more than one context

* A recreational context is characterized by a limited focus on competition, low intensity and commitment, formal organization but irregular and local involvement. A developmental context is characterized by a more formal competition structure, and the requirement for a greater commitment from players than exists in participatory sport. Players are also often selected through some form of talent identification. Finally, an elite context is characterized by intensive preparation and involvement from players, highly structured and formalized competition, and coaches who work with the same group of players in a full-time capacity (Trudel & Gilbert, 2006, p. 520-522).

Six studies observed coaches' behavior during matches/game, with eighteen during training. Only two studies observed coaches during both matches/game and training.

Table 10. Situations under which coaching behaviour was observed

Situation	Number of studies
Practice	18
Game	6
Both	2
Total	26

Discussion

The purpose of this paper was to review studies that had used a systematic observation method to investigate coaching behavior, and to consider the extent to which this area of research has developed since Kahan's (1999) review. The use of systematic observation to identify coaches' behavior has continued to receive substantial research interest and has undoubtedly provided important insights that have added to the body of sports coaching knowledge (Cushion, 2013). However, some of the problematic trends identified by Kahan (1999) still exist and will be overviewed in this section. The discussion will be presented under the four broader themes: 1) instrument development and technology, 2) Coder training, reliability, and procedural issues, 3) research questions and paradigm shift, 4) research context.

Instrument development and technology

Systematic observation instruments have been developed in line with advancements in technology. This has led to some movement away from instruments, which Kahan (1999) reported as being the most employed. These were, the Coach Behavior Recording Form (CBRF) (Langsdorf, 1979), the Coach Behavior Assessment System (CBAS) (Smith, Smoll, & Curtis, 1978) and most common, the Arizona State Observation Instrument (ASUOI) (Lacy & Darst, 1984). Since Kahan's review the CBRF and CBAS in particular have been employed sparingly by researchers undertaking systematic observation work, however, the ASUOI remains popular, although this has reduced as other systems have been validated and transposed onto a digital software platform. Several instruments have been developed based on existing instruments, which include the RCABI, the CAFIAS, and the

CAICS as it was claimed existing instruments did not enable the purpose of these studies to be met. Perhaps the most notable of the ‘newer’ systematic observation instruments, however, is the Coach Analysis and Intervention System (CAIS) (Cushion et al., 2012). It has been argued that the CAIS is a more sophisticated systematic observation method than those previous, as it provides a greater breakdown of coaching behaviors that better reflect those used by coaches, enables multi-level coding (i.e. coding more than one behavior at once), and allows researchers to code secondary behaviors (e.g. recipient, timing, content), and as a function of the practice form behaviors have occurred in (Cushion et al., 2012; Harvey et al., 2013).

Regardless of the systematic observation employed, there has been a trend in adapting or modifying the chosen instrument. For example, Bloom, Crumpton and Anderson (1999) and Zetou, Amprasi, Michalopoulou, and Aggelousis (2011) used a revised version of the CBRF, while Ford, Yates and Williams (2010) and Smith and Cushion (2006) among others used a modified version of the ASUOI. This suggests that these instruments were not appropriate in enabling researchers to gather data that satisfied their research questions. This could be explained by the dated nature of these instruments and their inability to reflect current thinking in coaching. However, a similar situation exists with the use of the CAIS. As evidenced in the results, a modified version of this instrument (i.e. Guzman & Calpe-Gomez, 2012; Partington & Cushion, 2012; Partington & Cushion, 2013) has been used more often than the validated version as presented by Cushion et al. (2012). Furthermore, only one study (Harvey et al., 2013) reported secondary behaviors. Finally, while researchers have made claims that they have used the CAIS, the primary behaviors were different to those noted by Cushion et al. (2012). It could be argued then that where this has occurred, the CAIS was in fact not the instrument employed.

Due to the range in systematic observation instruments used it is difficult to assess what coaches do in different sports and contexts. Different instruments include different behaviors, which are defined differently making it challenging for readers to interpret systematic observation data. For example, the RCABI as used by Hall, Gray and Sproule (2015) defined praise as: “*Non-specific praise given during the activity (e.g., “Excellent”, clapping)*” while the CAIS used by Harvey et al. (2013) defined the same behavior as: “*Positive or supportive verbal statements or non-verbal gestures 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". While it is appreciated that no one systematic observation method can be all encompassing and suit the purpose of every study, there is a critical need to use more common language when defining behaviors. Although not to advocate the use of one instrument over another, researchers do need to consider the instrument they are using and offer a suitable rationale for why they are employing this. For example, if a modified or adapted version of an instrument is being employed, why is this? Or, if a less sophisticated system is adopted over a system that is more complex, then what is the rationale for continuing with the less complex system that has been argued does not best capture coaches' behavior? From the studies reviewed that used modified or adapted versions of an instrument, researchers offered limited rationale of why the full version was not appropriate.

Research questions and paradigm shift

Early coaching research was conducted, interpreted and discussed through a positivistic lens (Gilbert & Trudel, 2004), as attempts were made to demonstrate a causal relationship between coach behavior and athlete response (Kahan, 1999). While the coach occupies a position of centrality and considerable influence on athletes' sporting performances (Cushion, 2010), it is now well appreciated that coaching is a social process with many factors influencing athlete learning (Cushion, 2013). It has been suggested that using systematic observation as an isolated method cannot appreciate the social contextual factors that can impact coaches' behaviors (Potrac, Jones, & Cushion, 2007). To investigate the socio-contextual elements of coaching, different research questions needed asking, which has resulted in the use of additional methods. Consequently, since Kahan's (1999) review this area of study has seen the emergence of mixed methodologies where qualitative methods have been used in conjunction with a systematic observation method.

The purpose of using qualitative methods, mainly in the form of interviews is that they enable researchers to gain an understanding of how and why coaches use certain behaviors and practice forms/activities (Smith & Cushion, 2006; Potrac, Jones, & Armour, 2002; Potrac et al., 2007; Partington, Cushion, & Harvey, 2014). Indeed, it has been suggested that to make changes to 'what' coaches do, there must be an understanding of 'why' they do it (Potrac et al., 2007). Interviews have been employed mostly with the coaches studied by the researchers, but in some cases with

coach's athletes (Webster, Hunt, & LeFleche, 2013) or with key stakeholders (i.e. parents) (Vinson et al., 2016) in order to investigate their perceptions of the coach's behavior. Another qualitative method used has been field note recordings in an attempt to examine coaching practice in greater detail (Stodter & Cushion, 2014; Vinson et al., 2016), however, this method has been used sparsely and is in need of greater research focus.

In attempts to understand why coaches use particular pedagogical strategies, scholars have drawn on, and introduced sociological theory and theoretical concepts and related these to coaching. For example, Potrac et al. (2002) interpreted their data through Goffman's (1959) concepts of 'social role' 'power' and 'presentation of the self', while Potrac et al. (2007) used French and Raven's work on power (1959) to offer explanations of why coaches used certain behaviors at the expense of others. Furthermore, Harvey et al. (2013) used Bruner's (1999) notion of 'folk pedagogies' when interpreting why the coaches in their study may have coached in particular ways. While this work is much welcomed and has offered a furthered understanding of coaches' practice, it seems the case that the theories drawn upon have not been well developed in coaching. In other words, scholars have tended to introduce many different theoretical concepts without, arguably, exploring these in any great depth. What appears needed is the development of existing theories and concepts used in coaching when theorizing practice, before the introduction of different theories. For example, the work of sociological theorists such as Erving Goffman has been used in studies where systematic observation has been the predominant data generation method (Potrac et al. 2002; Partington & Cushion, 2012), yet these have often been one-off studies conducted with a particular coach or group of coaches in a particular context. What we are advocating is that researchers to build upon this work and, thus, develop an increased understanding of how these theories can explain coaching practice.

Alongside a systematic observation method, other quantitative methods have been used. The most common method has been time-use analysis, which is a method that measures the amount of time a coach engages their athletes in different practice forms and activities. Although this method has been used previous to this review (i.e. Lacy & Martin, 1994) it has received more attention in recent years. Most studies that have used a time-use analysis method have examined the time coaches spend engaging athletes in 'training' or 'playing' form. These data have provided evidence

pertaining to how coaches are structuring their practice and whether they are engaging athletes in the most meaningful and relevant activities for their development. Harvey et al. (2013) went one step further than this and also recorded the time spent in 'other' form. This was any time when players were physically inactive. Findings from this study showed athletes spent considerable periods of time in this practice form. Given this, it does raise the question what periods of physical inactivity were coded as in other studies that used a time-use analysis method. As with issues related to coach behavior definitions, there is a need for greater consistency in how researchers are using systematic observation instruments and accompanying methods in order to gather data that are most reflective of a coach's practice.

Besides time-use analysis, other quantitative methods have been used. Becker and Wrisberg (2008) used a modified expectancy rating scale in order to measure whether coaches gave different types of feedback to athletes they regarded as either high expectancy (HE) or low expectancy (LE). As with the use of field notes, this method has been used sparingly, making specific conclusions and recommendations difficult in this current review.

The use of additional methods alongside systematic observation is a welcome development in systematic observation research, and something researchers should give serious consideration to when designing studies using systematic observation. These additional methods provide further insights into the nuances of the impact of the coaching context and how this implicates coaching behavior. Although systematic observation is considered one of the most appropriate means to identify what coaches do, coaches' behavior cannot be contextualized without a knowledge and understanding of why or how coaches employ certain behaviors (van der Mars, 1989; Cushion, 2010). This is important as it gives a sense of what coaches were trying to achieve and what factors informed their practice, and gives details pertaining to the interactions between coach and athlete (Groom, Nelson, & Cushion, 2012; Cope, Partington, Cushion, & Harvey, 2016), as well as other key stakeholders (i.e. administrators, parents).

Coder training, reliability, and procedural issues

Recently, Ayers and Blankenship (2015) conducted a presentation at the Physical Education Teacher Education conference in Atlanta titled, *Where Have All the Systematic Observation Instruments Gone?* While their main argument was based on

the reduction in utilization of these instruments in teacher education programs in the USA, the issue of training individuals in systematic observation procedures, and where these instruments appear in coach education and development programs, as well as in doctoral student programs is one worth considering. We would argue from our own experiences that this reduction in the utilization of these instruments in teacher education, and lack of use in coach education and development, may be due to there being a lack of researchers who clearly understand and are trained in behavioral analysis techniques during undergraduate, masters and doctoral level programs. This issue may be due to the fact that a range of methodologies to examine coaching practice has developed, as argued in the previous section. However, we would argue that to gain an in-depth understanding of what coaches do, and how this changes over time, this requires some form of behavioral analysis assessment (Cushion et al., 2012). Consequently, this raises additional issues about offering quality coder training, where this appears in coach education and development and in doctoral programs, as well as the need to follow strict training procedures.

Ensuring the credibility of data is essential when employing a systematic observation method (McKenzie & van der Mars, 2015). McKenzie and van der Mars (2015) consider data can only be credible if coders have been through a process of proper training, and reliability checks are conducted throughout data collection and analysis. Indeed, McKenzie and van der Mars (2015) offer a coding training protocol to follow, however, while seventeen of the twenty-six studies stated that coders had been trained, it is unclear from this review the extent to which studies have followed this coder training protocol, or something similar. As such, data presented where coder training and reliability has not been reported should be read with caution as there are no means of detecting whether these data are representative of what coaches actually do.

Prolonged observations of coaches during the different phases of a season are another mechanism by which to ensure data are representative of coaches' behavior. Kahan (1999) was highly critical of a single observation of coaches suggesting that conclusions could not be drawn from such 'snapshots'. Unfortunately, similar issues exist despite continued calls for work of a more season-long and/or longitudinal nature (Harvey et al., 2013; Kahan, 1999). The general pattern is that studies with smaller sample sizes often observe coaches for longer, and those with increased sample sizes conduct fewer observations. As Kahan (1999) acknowledged,

researchers face a decision of whether they choose a larger sample size and thus limit the number of observations, or choose a smaller sample size and increase the number of observations. There comes a point, however, when a minimum number of observations are required if data are to be representative of what coaches do, which Brewer and Jones (2002) suggest to be three coaching sessions of 90 minutes per coach. The problem with single observations, or limited time spent observing is that coaches may act or behave in certain ways to satisfy the observation period (Partington & Cushion, 2012). Equally, due to the contextual and situational nature of coaching, a single observation cannot be deemed an example of how a coach behaves, and should be avoided.

Another issue is that without long periods of time in the field conducting observations, it is impossible to undertake intervention-based studies. With the exception of studies by Stodter and Cushion (2014) and Partington, Cushion, Cope and Harvey (2015), no other studies in the current review investigated *changes* in coaches' behavior, which means that little is known about how to most effectively do this (More & Franks, 1996). Therefore, while descriptive examinations of practice provide information of what coaches are doing and are therefore essential (Potrac et al., 2007), if an understanding is to be developed regarding the impact of different learning interventions on coaches' behavior and practice, then intervention studies are a necessity.

Given the time consuming nature of collecting and analyzing systematic observation data, it is unsurprising that few studies have moved beyond a small number of observations, or carried out seasonal or longitudinal interventions. This issue has somewhat not been helped by the introduction of more sophisticated and complex systematic instruments. Consequently, while these systems are welcomed for providing a greater level of information regarding a coach's behavior and practice, the tradeoff is that coder training, analysis of data, and achieving the required level of reliability is a more onerous, and challenging process. Yet, if researchers want to investigate such things as how coaches' behavior changes over the course of season, or during different phases of a season, more seasonal/longitudinal work is required using a systematic observation instrument that appropriately captures what coaches are doing, rather than utilizing a system that is perhaps easier to use and more convenient.

The depth to which systematic observation data can be collected is dependent on whether live, or post-observation coding takes place. Coinciding with developments in systematic observation instruments is the use of video to record coaching sessions. Although there are advantages to this, such as the ability to code primary and secondary behaviors and conduct post-observation reliability tests, there are also feasibility issues that need consideration. For example, the more complex the instrument the more challenging it is to reliably capture all information as the coaching is happening. As such, if researchers wanted to use a system such as CAIS, they would have no choice but to code post event, unless they used a modified or adapted version like Partington and colleagues (2012; 2013; 2014).

Coding post event opens up the possibility to use a time-sampled event method, which is coding each behavior every time it occurs (van der Mars, 1989). However, the method of coding researchers decide to employ is not so much the issue as them offering an explanation of what they mean by this method of coding. In most cases, researchers state what method of coding they have employed but fail to tell readers what this method is. Researchers need to address this issue and offer greater clarity over the method of coding used.

A final issue with the use of video is that there are increased ethical constraints that require consideration. While this should not be a determining factor in researchers using video, pragmatically it could be problematic when observing coaches of children, and in certain sports (i.e. swimming and gymnastics). Also, if the coaching takes place in an area accessed by other people as most recreational children's sessions do this leads to further ethical issues of making sure anyone who could appear in the video is aware and consenting. While this is a challenge, this could be overcome by making clear to participants that techniques such as pixelating faces and clothing will help ensure anonymity of identity.

Research Context

Where systematic observation research in coaching has developed is through systematically observing coaches across different contextual domains (e.g. participation, development and performance). Kahan (1999) reported that during the period of his review there had been a predominant focus on studies that had systematically observed the behaviors of coaches in youth sport contexts. Although youth sport stretches across the participant and development domains, this current

review highlighted that more research was being conducted in the performance domain. This is a positive sign and is something researchers need to continue doing if detailed understandings of what coaches do in different contextual domains is to be gained.

We did identify two primary gaps in the current literature on systematic observation through our analysis. First, the systematic observation of coaches during training, rather than games continues to dominate systematic observation research. Although there are fewer games than practices, coaches' behaviours have been identified as being different under the two conditions, hence practice behaviours can not be assumed as being the same as coaches in-game behaviours (Cushion, 2010; Trudel et al., 1996). Therefore, a complete picture of coaching behaviour and the potential relationships between practice and in-competition behaviours has not been examined, especially in the range of contexts where coaching takes place. Certainly, coaches have much more control over their behaviour within their own practice environments. In contrast, coaches' behaviour may be more reactionary in competition, where coaches make decisions in response to the continually changing environment, and more circumstances that are beyond their own control.

Second, following on from an argument made by Kahan (1999) there remains a limited understanding of coaches' behavior across a variety of sports, countries and coaching populations. Although it could be claimed that there is much systematic observation data to draw on to provide evidence of what coaches do, it is the case that these data have been generated mainly from male coaches, in a limited number of sports, and in a select number of countries. We similarly found it is mostly the same sports that are still receiving the majority of the research attention in this area. This is not to criticize the research we reviewed, as it has been most helpful in providing an in-depth understanding of what coaches in these sports are doing. However, a number of 'gaps' remain in systematic observation research with respect to context. For example, research with female coaches at all levels is of critical need. Moreover, the behavior and practice of coaches who work in a disability domain is urgently needed, as well as research about the role that assistant coaches play in training and competition games (Gilbert & Trudel, 2004; Hall et al., 2015).

Conclusion

This review has shown that while systematic observation continues to advance knowledge and understanding of what coaches do, there are many areas, as highlighted in the discussion section that require further research attention. Without wishing to repeat these here, we do urge researchers to adopt a more critical approach when adopting a systematic observation method. This includes researchers offering a clearer rationale for the systematic observation instrument being employed, considering the number of observations for each coach, and reflecting on the use of a multiple, mixed methods approach. We hope that this review has brought some of these issues to light, and offers greater clarity for researchers and practitioners wanting to employ this method in future work. Furthermore, we hope this review acts as a useful guide for editors and reviewers who are responsible for making judgments of the quality of this type of work.

References

- Ayers, S. F., Blankenship, B. T. (2015, October). *Where have all the systematic observation tools gone?* Physical Education in Higher Education Conference, Atlanta, GA, USA.
- *Becker, A. J., & Wrisberg, C. A. (2008). Effective coaching in action: Observations of legendary collegiate basketball coach Pat Summitt. *Sport Psychologist*, 22(2), 197-211.
- *Bloom, G. A., Crumpton, R., & Anderson, J. E. (1999). A systematic observation study of the teaching behaviors of an expert basketball coach. *Sport Psychologist*, 13, 157-170.
- Brewer, C. J., & Jones, R. L. (2002). A five-stage process for establishing contextually valid systematic observation instruments: the case of rugby union. *Sport Psychologist*, 16(2), 138-159.
- Cope, E., Partington, M., Cushion, C. J., & Harvey, S. (2016). An investigation of professional top-level youth football coaches' questioning practice. *Qualitative Research in Sport, Exercise and Health*, 1-14.
- Cushion, C.J. (2010). Coach behavior. In J. Lyle & C.J. Cushion (Eds.), *Sports Coaching Professionalization and Practice* (pp. 43-62). London: Elsevier.
- Cushion, C. J. (2013). Applying game centered approaches in coaching: a critical analysis of the 'dilemmas of practice' impacting change. *Sports Coaching Review*, 2(1), 61-76.
- *Cushion, C. J., & Jones, R. L. (2001). A systematic observation of professional top-level youth soccer coaches. *Journal of Sport Behaviour*, 24(4), 354-376.
- Cushion, C., Harvey, S., Muir, B., & Nelson, L. (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.
- *Erickson, K., Côté, J., Hollenstein, T., & Deakin, J. (2011). Examining coach-athlete interactions using state space grids: An observational analysis in competitive youth sport. *Psychology of Sport and Exercise*, 12(6), 645-654.
- *Ford, P. R., Yates, I., & Williams, A. M. (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(5), 483-495.

- Gilbert, W. D., & Trudel, P. (2004). Analysis of coaching science research published from 1970–2001. *Research quarterly for exercise and sport*, 75(4), 388-399.
- *Gomez, V.C., Guzman Lujan, J.F. & Santamaria, C.G. (2013) Relationship between score and coaches' verbal behaviour. *Journal of Human Sport & Exercise*. 8 Proc. 3 vol: 19. 728-737.
- Groom, R., Cushion, C. J., & Nelson, L. J. (2012). Analysing coach–athlete 'talk in interaction' within the delivery of video-based performance feedback in elite youth soccer. *Qualitative research in sport, exercise and health*, 4(3), 439-458.
- *Guzmán Luján, J. F., & Calpe Gómez, V. (2012). Preliminary study of coach verbal behaviour according to game actions. *Journal of Human Sport and Exercise*, 7(2), 376-382.
- *Harvey, S., Cushion, C. J., Cope, E., & Muir, B. (2013). A season long investigation into coaching behaviours as a function of practice state: The case of three collegiate coaches. *Sports Coaching Review*, 2(1), 13-32.
- Harvey, S., & Jarrett, K. (2014). A review of the game-centred approaches to teaching and coaching literature since 2006. *Physical Education and Sport Pedagogy*, 19(3), 278-300.
- *Hall, E. T., Gray, S., & Sproule, J. (2015). The microstructure of coaching practice: behaviours and activities of an elite rugby union head coach during preparation and competition. *Journal of sports sciences*, 1-10.
- *Horton, S., Baker, J., & Deakin, J. (2005). Expert in action: a systematic observation of 5 national team coaches. *International Journal of Sport Psychology*, 36(4), 299
- *Isabel, M., Antonio, S., Antonio, R., Rosado, F., and Michel, M. M. (2008). A Systematic Observation of Youth Amateur Volleyball Coaches Behavior, *International Journal of Applied Sports Science*, 10 (2), pp. 37-59.
- Kahan, D. (1999). Coaching behavior: A review of the systematic observation research literature. *Applied Research in Coaching and Athletics Annual*, 17-58.
- Lacy, A. C., & Darst, P. W. (1984). Evolution of a systematic observation system: The ASU coaching observation instrument. *Journal of Teaching in Physical Education*, 3(3), 59-66.

- Lacy, A. C., & Martin, D. L. (1994). Analysis of starter/nonstarter motor-skill engagement and coaching behaviors in collegiate women's volleyball. *Journal of Teaching in Physical Education*, 13, 95-95.
- Langsdorf, E. V. (1979). *A systematic observation of football coaching behavior in a major university environment*. Unpublished Doctoral Thesis, Arizona State University, Tempe.
- LaVoi, N. M., & Dutove, J. K. (2012). Barriers and supports for female coaches: An ecological model. *Sports Coaching Review*, 1(1), 17-37.
- *Lorenzo, Navarro, Rivilla & Lorenzo, A. (2013) The analysis of basketball coach speech during the moments of game and pause in relation to the performance in completion. *Revista de Psicología del Deporte*. 22 (1) 227-230.
- Lyle, J. (2002). *Sports coaching concepts: A framework for coaches' behaviour*. London: Routledge.
- *Massey, C. D., Maneval, M. W., Phillips, J., Vincent, J., White, G., & Zoeller, B. (2002). An analysis of teaching and coaching behaviors of elite strength and conditioning coaches. *The Journal of Strength & Conditioning Research*, 16(3), 456-460
- McKenzie, T. L., & Van Der Mars, H. (2015). Top 10 research questions related to assessing physical activity and its contexts using systematic observation. *Research quarterly for exercise and sport*, 86(1), 13-29.
- More, K. G., & Franks, I. M. (1996). Analysis and modification of verbal coaching behaviour: The usefulness of a data- driven intervention strategy. *Journal of Sports Sciences*, 14(6), 523-543.
- *Partington, M., & Cushion, C. J. (2012). Performance during performance: Using Goffman to understand the behaviours of elite youth football coaches during games. *Sports Coaching Review*, 1(2), 93-105.
- *Partington, M., & Cushion, C. (2013). An investigation of the practice activities and coaching behaviors of professional top- level youth soccer coaches. *Scandinavian journal of medicine & science in sports*, 23(3), 374-382.
- *Partington, M., Cushion, C., & Harvey, S. (2014). An investigation of the effect of athletes' age on the coaching behaviours of professional top-level youth soccer coaches. *Journal of sports sciences*, 32(5), 403-414.
- *Partington, M., Cushion, C. J., Cope, E., & Harvey, S. (2015). The impact of video feedback on professional youth football coaches' reflection and practice

- behaviour: a longitudinal investigation of behaviour change. *Reflective Practice*, 16(5), 700-716.
- *Pereira, F., Mesquita, I., & Graca, A. (2009). Accountability systems and instructional approaches in youth Volleyball training. *Journal of Sports Sciences & Medicine*, 8(3), 366.
- *Pereira, F. R. M., Mesquita, I. M. R., & Graça, A. B. (2010). Relating content and nature of information when teaching volleyball in youth volleyball training settings. *Kineziologija*, 42(2), 121-131.
- *Potrac, P., Jones, R., & Armour, K. (2002). 'It's All About Getting Respect': The Coaching Behaviors of an Expert English Soccer Coach. *Sport, Education and Society*, 7(2), 183-202.
- *Potrac, P., Jones, R., & Cushion, C. (2007). Understanding power and the coach's role in professional English soccer: A preliminary investigation of coach behaviour. *Soccer and Society*, 8(1), 33-49.
- *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., & Curtis, B. (1978). Coaching behaviors in little league baseball. *Psychological perspectives in youth sports*, 173, 201-208.
- *Stodter, A., & Cushion, C. J. (2014). Coaches' learning and education: a case study of cultures in conflict. *Sports Coaching Review*, 3(1), 63-79.
- Trudel, P., Côté, J., & Bernard, D. (1996). Systematic observation of youth ice hockey coaches during games. *Journal of Sport Behavior*, 19(1), 50-65
- Trudel, P., & Gilbert, W. (2006). Coaching and coach education. In D. Kirk. D. MacDonald, & M. O'Sullivan (Eds.), *Handbook of physical education* (pp. 516-539). London: Sage Publications.
- Van der Mars, H. (1989). Observer reliability: Issues and procedures. In P. W. Darst. D. Zakrajsek, & V. H. Mancini. *Analyzing Physical Education and Sport Instruction* (Eds.), (2nd edn.) (pp. 53-80). Champaign, IL: Human Kinetics.
- *Vinson, D., Brady, A., Moreland, B., & Judge, N. (2016). Exploring Coach Behaviours, Session Contexts and Key Stakeholder Perceptions of Non-linear Coaching Approaches in Youth Sport. *International Journal of Sports Science and Coaching*, 11(1).

- *Webster, C. A., Hunt, K., & LaFleche, M. (2013). Are Winning Coaches More Autonomy-Supportive? Examining the Context of Varsity Boys' Soccer. *Journal of Sport Behavior*, 36(2), 209.
- *Zetou, E., Amprasi, E., Michalopoulou, M., & Aggelousis, N. (2011). Volleyball coaches behavior assessment through systematic observation. *Journal of Human Sport and Exercise Science*, 6(4), 585-593.

Note: References were an asterisk signify those papers included in the review