

Cope, E., Partington, M. & Harvey, S. (2016) A review of the use of a systematic observation method in coaching research between 1975-2016, *Journal of Sports Sciences*. DOI: 10.1080/02640414.2016.1252463

1 **Abstract**

2
3 A systematic observation method has been one of the most popularly employed
4 methods in coaching research. Kahan's review of this method conducted between
5 1975-1997 highlighted the key trends in this research, and offered methodological
6 guidance for researchers wishing to use this method in their research. The purpose of
7 this review was to provide an update of the use of a systematic observation method in
8 coaching research and assess the extent to which the calls made by Kahan have been
9 addressed. While in some respect this field of study has progressed (i.e. the
10 introduction of qualitative methods), researchers adopting this method have failed to
11 attend to many of the issues Kahan raised. For this method to continue to make a
12 positive contribution toward the coaching research literature, researchers need to more
13 critically reflect on how and why they are employing this method. At present, some of
14 the decisions made by researchers who have conducted work in this area are not
15 justified with a rationale. It is our intention that this review will serve as guidance for
16 researchers and practitioners, and editors and reviewers of journals when attempting
17 to assess the quality of this type of work.

18 **Introduction**

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

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

42 In the period since Kahan's review, the use of systematic observation has
43 remained popular amongst coaching researchers, and has continued to evolve as the
44 field of coaching has become more established. However, the extent to which this
45 evidence has contributed toward the development of coaching practice, especially
46 within the confines of the specific contexts in which 'coaching' takes place (Lyle,
47 2002), is unclear. Therefore, the purpose of this paper is to present an updated review
48 of research into the use of a systematic observation method to record coaching
49 behavior, and consider how this line of research has moved forward since Kahan's
50 review. First, we overview how we identified studies to be included in the review.

51

52 **Method**

53 *Identification of studies*

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

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

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

86

87 ***Summary of Studies from 1997 to 2016***

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

101

102 **Results**

103 ***Instrument development and technology***

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

120 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

121 Note: The total equals 29 because some studies employed more than one systematic observation method.
 122

123 *Coder training, reliability, and procedural issues*

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

131

132 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)

133 Note: The total equals 25 because one study employed a consensus building technique
 134

135 There were differences in the number of coaches observed for each study depending
 136 on its purpose and nature. However, not all coaches were observed within the same
 137 study the same number of times. For example, in Harvey, Cushion, Cope and Muir
 138 (2013) study, the three coaches were observed a different number of times each. In
 139 these instances we grouped coaches together and reported the mean number of
 140 observations per study. In six studies, coaches were observed only once, in seven
 141 studies coaches were observed for an average of between two to four times, in six
 142 studies coaches were observed for an average of between five to seven times, in three
 143 studies coaches were observed for an average of between eight to ten times, and in
 144 only one study a coach was observed on more than ten occasions. Two studies did not
 145 report the number of times coaches were observed, while one study was incalculable.

146

147 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

148 Note: Incalculable: studies in which a set datum value across coaches was reported for observation frequency.
149

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

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

163

164 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

165 Note: This number equals 29 because some studies used two methods of recording
166

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

169

170 Table 5. Method of systematic observation recording

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

171

172 ***Research questions and paradigm shift***

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

184

185 Table 6. Additional methods implemented

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

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

187

188 ***Research Context***

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

194

195 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

196 Note: The total equals 30 because some studies used a systematic observation method to investigate coaches'
 197 behaviour in more than one sport
 198

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

204

205 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

206

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

210

211 Table 9. Coaching domain reported across studies

Domain	Frequency
Participation	6
Development	12
Performance	12
Total	30

212 Note: This total equals 30 because some studied used a systematic observation method across more than one
 213 context
 214

* 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).

215

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

219

220 Table 10. Situations under which coaching behaviour was observed

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

221

222 Discussion

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

233

234 *Instrument development and technology*

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

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

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

272 Due to the range in systematic observation instruments used it is difficult to
273 assess what coaches do in different sports and contexts. Different instruments include
274 different behaviors, which are defined differently making it challenging for readers to
275 interpret systematic observation data. For example, the RCABI as used by Hall, Gray
276 and Sproule (2015) defined praise as: “*Non-specific praise given during the activity*
277 *(e.g., “Excellent”, clapping)”* while the CAIS used by Harvey et al. (2013) defined
278 the same behavior as: “*Positive or supportive verbal statements or non-verbal*
279 *gestures which demonstrate the coach’s general satisfaction or pleasure to a*

280 *player(s) that DO NOT specifically aim to improve the player(s) performance at the*
281 *next skill attempt*". While it is appreciated that no one systematic observation method
282 can be all encompassing and suit the purpose of every study, there is a critical need to
283 use more common language when defining behaviors. Although not to advocate the
284 use of one instrument over another, researchers do need to consider the instrument
285 they are using and offer a suitable rationale for why they are employing this. For
286 example, if a modified or adapted version of an instrument is being employed, why is
287 this? Or, if a less sophisticated system is adopted over a system that is more complex,
288 then what is the rationale for continuing with the less complex system that has been
289 argued does not best capture coaches' behavior? From the studies reviewed that used
290 modified or adapted versions of an instrument, researchers offered limited rationale of
291 why the full version was not appropriate.

292

293 ***Research questions and paradigm shift***

294 Early coaching research was conducted, interpreted and discussed through a
295 positivistic lens (Gilbert & Trudel, 2004), as attempts were made to demonstrate a
296 causal relationship between coach behavior and athlete response (Kahan, 1999).
297 While the coach occupies a position of centrality and considerable influence on
298 athletes' sporting performances (Cushion, 2010), it is now well appreciated that
299 coaching is a social process with many factors influencing athlete learning (Cushion,
300 2013). It has been suggested that using systematic observation as an isolated method
301 cannot appreciate the social contextual factors that can impact coaches' behaviors
302 (Potrac, Jones, & Cushion, 2007). To investigate the socio-contextual elements of
303 coaching, different research questions needed asking, which has resulted in the use of
304 additional methods. Consequently, since Kahan's (1999) review this area of study has
305 seen the emergence of mixed methodologies where qualitative methods have been
306 used in conjunction with a systematic observation method.

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

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

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

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

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

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

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

377

378 ***Coder training, reliability, and procedural issues***

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

382 the reduction in utilization of these instruments in teacher education programs in the
383 USA, the issue of training individuals in systematic observation procedures, and
384 where these instruments appear in coach education and development programs, as
385 well as in doctoral student programs is one worth considering. We would argue from
386 our own experiences that this reduction in the utilization of these instruments in
387 teacher education, and lack of use in coach education and development, may be due to
388 there being a lack of researchers who clearly understand and are trained in behavioral
389 analysis techniques during undergraduate, masters and doctoral level programs. This
390 issue may be due to the fact that a range of methodologies to examine coaching
391 practice has developed, as argued in the previous section. However, we would argue
392 that to gain an in-depth understanding of what coaches do, and how this changes over
393 time, this requires some form of behavioral analysis assessment (Cushion et al.,
394 2012). Consequently, this raises additional issues about offering quality coder
395 training, where this appears in coach education and development and in doctoral
396 programs, as well as the need to follow strict training procedures.

397 Ensuring the credibility of data is essential when employing a systematic
398 observation method (McKenzie & van der Mars, 2015). McKenzie and van der Mars
399 (2015) consider data can only be credible if coders have been through a process of
400 proper training, and reliability checks are conducted throughout data collection and
401 analysis. Indeed, McKenzie and van der Mars (2015) offer a coding training protocol
402 to follow, however, while seventeen of the twenty-six studies stated that coders had
403 been trained, it is unclear from this review the extent to which studies have followed
404 this coder training protocol, or something similar. As such, data presented where
405 coder training and reliability has not been reported should be read with caution as
406 there are no means of detecting whether these data are representative of what coaches
407 actually do.

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

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

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

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

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

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

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

475

476 ***Research Context***

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

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

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

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

514

515 **Conclusion**

516 This review has shown that while systematic observation continues to advance
517 knowledge and understanding of what coaches do, there are many areas, as
518 highlighted in the discussion section that require further research attention. Without
519 wishing to repeat these here, we do urge researchers to adopt a more critical approach
520 when adopting a systematic observation method. This includes researchers offering a
521 clearer rationale for the systematic observation instrument being employed,
522 considering the number of observations for each coach, and reflecting on the use of a
523 multiple, mixed methods approach. We hope that this review has brought some of
524 these issues to light, and offers greater clarity for researchers and practitioners
525 wanting to employ this method in future work. Furthermore, we hope this review acts
526 as a useful guide for editors and reviewers who are responsible for making judgments
527 of the quality of this type of work.
528

529 **References**

- 530 Ayers, S. F., Blankenship, B. T. (2015, October). *Where have all the systematic*
531 *observation tools gone?* Physical Education in Higher Education Conference,
532 Atlanta, GA, USA.
- 533 *Becker, A. J., & Wrisberg, C. A. (2008). Effective coaching in action: Observations
534 of legendary collegiate basketball coach Pat Summitt. *Sport Psychologist*,
535 22(2), 197-211.
- 536 *Bloom, G. A., Crumpton, R., & Anderson, J. E. (1999). A systematic observation
537 study of the teaching behaviors of an expert basketball coach. *Sport*
538 *Psychologist*, 13, 157-170.
- 539 Brewer, C. J., & Jones, R. L. (2002). A five-stage process for establishing
540 contextually valid systematic observation instruments: the case of rugby
541 union. *Sport Psychologist*, 16(2), 138-159.
- 542 Cope, E., Partington, M., Cushion, C. J., & Harvey, S. (2016). An investigation of
543 professional top-level youth football coaches' questioning practice.
544 *Qualitative Research in Sport, Exercise and Health*, 1-14.
- 545 Cushion, C.J. (2010). Coach behavior. In J. Lyle & C.J. Cushion (Eds.), *Sports*
546 *Coaching Professionalization and Practice* (pp. 43-62). London: Elsevier.
- 547 Cushion, C. J. (2013). Applying game centered approaches in coaching: a critical
548 analysis of the 'dilemmas of practice' impacting change. *Sports Coaching*
549 *Review*, 2(1), 61-76.
- 550 *Cushion, C. J., & Jones, R. L. (2001). A systematic observation of professional top-
551 level youth soccer coaches. *Journal of Sport Behaviour*, 24(4), 354-376.
- 552 Cushion, C., Harvey, S., Muir, B., & Nelson, L. (2012). Developing the Coach
553 Analysis and Intervention System (CAIS): Establishing validity and reliability
554 of a computerised systematic observation instrument. *Journal of sports*
555 *sciences*, 30(2), 201-216.
- 556 *Erickson, K., Côté, J., Hollenstein, T., & Deakin, J. (2011). Examining coach-
557 athlete interactions using state space grids: An observational analysis in
558 competitive youth sport. *Psychology of Sport and Exercise*, 12(6), 645-654.
- 559 *Ford, P. R., Yates, I., & Williams, A. M. (2010). An analysis of practice activities
560 and instructional behaviours used by youth soccer coaches during practice:
561 Exploring the link between science and application. *Journal of Sports*
562 *Sciences*, 28(5), 483-495.

- 563 Gilbert, W. D., & Trudel, P. (2004). Analysis of coaching science research published
564 from 1970–2001. *Research quarterly for exercise and sport*, 75(4), 388-399.
- 565 *Gomez, V.C., Guzman Lujan, J.F. & Santamaria, C.G. (2013) Relationship between
566 score and coaches' verbal behaviour. *Journal of Human Sport & Exercise*. 8
567 Proc. 3 vol: 19. 728-737.
- 568 Groom, R., Cushion, C. J., & Nelson, L. J. (2012). Analysing coach–athlete 'talk in
569 interaction' within the delivery of video-based performance feedback in elite
570 youth soccer. *Qualitative research in sport, exercise and health*, 4(3), 439-
571 458.
- 572 *Guzmán Luján, J. F., & Calpe Gómez, V. (2012). Preliminary study of coach verbal
573 behaviour according to game actions. *Journal of Human Sport and Exercise*,
574 7(2), 376-382.
- 575 *Harvey, S., Cushion, C. J., Cope, E., & Muir, B. (2013). A season long investigation
576 into coaching behaviours as a function of practice state: The case of three
577 collegiate coaches. *Sports Coaching Review*, 2(1), 13-32.
- 578 Harvey, S., & Jarrett, K. (2014). A review of the game-centred approaches to teaching
579 and coaching literature since 2006. *Physical Education and Sport Pedagogy*,
580 19(3), 278-300.
- 581 *Hall, E. T., Gray, S., & Sproule, J. (2015). The microstructure of coaching practice:
582 behaviours and activities of an elite rugby union head coach during
583 preparation and competition. *Journal of sports sciences*, 1-10.
- 584 *Horton, S., Baker, J., & Deakin, J. (2005). Expert in action: a systematic observation
585 of 5 national team coaches. *International Journal of Sport Psychology*, 36(4),
586 299
- 587 *Isabel, M., Antonio, S., Antonio, R., Rosado, F., and Michel, M. M. (2008). A
588 Systematic Observation of Youth Amateur Volleyball Coaches Behavior,
589 *International Journal of Applied Sports Science*, 10 (2), pp. 37-59.
- 590 Kahan, D. (1999). Coaching behavior: A review of the systematic observation
591 research literature. *Applied Research in Coaching and Athletics Annual*, 17-
592 58.
- 593 Lacy, A. C., & Darst, P. W. (1984). Evolution of a systematic observation system:
594 The ASU coaching observation instrument. *Journal of Teaching in Physical
595 Education*, 3(3), 59-66.

- 596 Lacy, A. C., & Martin, D. L. (1994). Analysis of starter/nonstarter motor-skill
597 engagement and coaching behaviors in collegiate women's volleyball. *Journal*
598 *of Teaching in Physical Education*, 13, 95-95.
- 599 Langsdorf, E. V. (1979). *A systematic observation of football coaching behavior in a*
600 *major university environment*. Unpublished Doctoral Thesis, Arizona State
601 University, Tempe.
- 602 LaVoi, N. M., & Dutove, J. K. (2012). Barriers and supports for female coaches: An
603 ecological model. *Sports Coaching Review*, 1(1), 17-37.
- 604 *Lorenzo, Navarro, Rivilla & Lorenzo, A. (2013) The analysis of basketball coach
605 speech during the moments of game and pause in relation to the performance
606 in completion. *Revista de Psicología del Deporte*. 22 (1) 227-230.
- 607 Lyle, J. (2002). *Sports coaching concepts: A framework for coaches' behaviour*.
608 London: Routledge.
- 609 *Massey, C. D., Maneval, M. W., Phillips, J., Vincent, J., White, G., & Zoeller, B.
610 (2002). An analysis of teaching and coaching behaviors of elite strength and
611 conditioning coaches. *The Journal of Strength & Conditioning Research*,
612 16(3), 456-460
- 613 McKenzie, T. L., & Van Der Mars, H. (2015). Top 10 research questions related to
614 assessing physical activity and its contexts using systematic observation.
615 *Research quarterly for exercise and sport*, 86(1), 13-29.
- 616 More, K. G., & Franks, I. M. (1996). Analysis and modification of verbal coaching
617 behaviour: The usefulness of a data-driven intervention strategy. *Journal of*
618 *Sports Sciences*, 14(6), 523-543.
- 619 *Partington, M., & Cushion, C. J. (2012). Performance during performance: Using
620 Goffman to understand the behaviours of elite youth football coaches during
621 games. *Sports Coaching Review*, 1(2), 93-105.
- 622 *Partington, M., & Cushion, C. (2013). An investigation of the practice activities and
623 coaching behaviors of professional top-level youth soccer coaches.
624 *Scandinavian journal of medicine & science in sports*, 23(3), 374-382.
- 625 *Partington, M., Cushion, C., & Harvey, S. (2014). An investigation of the effect of
626 athletes' age on the coaching behaviours of professional top-level youth
627 soccer coaches. *Journal of sports sciences*, 32(5), 403-414.
- 628 *Partington, M., Cushion, C. J., Cope, E., & Harvey, S. (2015). The impact of video
629 feedback on professional youth football coaches' reflection and practice

630 behaviour: a longitudinal investigation of behaviour change. *Reflective*
631 *Practice*, 16(5), 700-716.

632 *Pereira, F., Mesquita, I., & Graca, A. (2009). Accountability systems and
633 instructional approaches in youth Volleyball training. *Journal of Sports*
634 *Sciences & Medicine*, 8(3), 366.

635 *Pereira, F. R. M., Mesquita, I. M. R., & Graça, A. B. (2010). Relating content and
636 nature of information when teaching volleyball in youth volleyball training
637 settings. *Kineziologija*, 42(2), 121-131.

638 *Potrac, P., Jones, R., & Armour, K. (2002). 'It's All About Getting Respect': The
639 Coaching Behaviors of an Expert English Soccer Coach. *Sport, Education and*
640 *Society*, 7(2), 183-202.

641 *Potrac, P., Jones, R., & Cushion, C. (2007). Understanding power and the coach's
642 role in professional English soccer: A preliminary investigation of coach
643 behaviour. *Soccer and Society*, 8(1), 33-49.

644 *Smith, M., & Cushion, C. J. (2006). An investigation of the in-game behaviours of
645 professional, top-level youth soccer coaches. *Journal of sports sciences*, 24(4),
646 355-366.

647 Smith, R. E., Smoll, F. L., & Curtis, B. (1978). Coaching behaviors in little league
648 baseball. *Psychological perspectives in youth sports*, 173, 201-208.

649 *Stodter, A., & Cushion, C. J. (2014). Coaches' learning and education: a case study
650 of cultures in conflict. *Sports Coaching Review*, 3(1), 63-79.

651 Trudel, P., Côté, J., & Bernard, D. (1996). Systematic observation of youth ice
652 hockey coaches during games. *Journal of Sport Behavior*, 19(1), 50-65

653 Trudel, P., & Gilbert, W. (2006). Coaching and coach education. In D. Kirk, D.
654 MacDonald, & M. O'Sullivan (Eds.), *Handbook of physical education* (pp.
655 516-539). London: Sage Publications.

656 Van der Mars, H. (1989). Observer reliability: Issues and procedures. In P. W. Darst,
657 D. Zakrajsek, & V. H. Mancini. *Analyzing Physical Education and Sport*
658 *Instruction* (Eds.), (2nd edn.) (pp. 53-80). Champaign, IL: Human Kinetics.

659 *Vinson, D., Brady, A., Moreland, B., & Judge, N. (2016). Exploring Coach
660 Behaviours, Session Contexts and Key Stakeholder Perceptions of Non-linear
661 Coaching Approaches in Youth Sport. *International Journal of Sports Science*
662 *and Coaching*, 11(1).

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

669

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

671