National Analysis and Reliability Testing in Soccer: A Mobile Phone Application Approach

Rabiu Muazu Musa[1], Mohammad Razali Abdullah[1], Ahmad Bisyri Husin Musawi Maliki[1], Siti Musliha Mat-Rasid[2], Norlaila Azura Kosni[1], Aleesha Adnan[1]

[1]Faculty of Applied Social Science, Universiti Sultan Zainal Abidin, 21300 Gong Badak, Terengganu, Malaysia.
[2]East Coast Environmental Research Institute (ESERI), Universiti Sultan Zainal Abidin, 21300 Gong Badak, Terengganu, Malaysia
Email: razali896@yahoo.com

Abstract
Delivery of accurate information on the players' performance in soccer is considered to be vital in enhancing performance and overall coaching process. The provision of such accurate information can be promising when the application used is valid and reliable. A chartered sports engineer developed a cost-effective Mobile phone application called *StatWatch* for the analysis of players' performance both during training and competitions. This study aims to investigate the reliability of the application in the delivery of information on the players' performance. Eleven performance analysts with a mean age of 24.62 and ± 3.5 years' experience were required to notate the performance of one of the professional soccer team contending in a Malaysian Super League for four competitive matches consecutively via the application. The intra-class correlation coefficient was used to ascertain the level of the agreement of the performance analysts based on the data recorded. The intra-class average measure reveals 0.98 which explains that 98% of the variance and the means of the analysts in their overall analysis is real $p < 0.001$. Similarly, a Cronbach's Alpha co-efficient indicates .099 which suggests an excellent consistency in their general observation. An Android based application is found to be reliable in the notational analysis of soccer. Professional and semi-professional soccer clubs could consider Android based application in analysing their players' performance due to its effectiveness in proving information on performance.
Keywords*: Mobile phone application, Notational analysis, Reliability testing, Soccer.

Introduction
Performance analysis is a part of sports and exercise science that measures the real performance of an athlete as instead of views or self-report by the athlete, coach or an observer (O'donoghue, 2010). Performance analysis is primarily utilised in numerous sports

to enable coaches to acquire an objective data that can be used to give feedback on performance. Providing feedback on performance in soccer is reported to play a significant role in improving the overall performance of a player (Abdullah et al., 2016a; Hook & Hughes 2001; Sermento et al., 2014). Performance analysis in soccer is widely recognised by many professional soccer clubs in the world to provide objective information that can improve team performance. Professional soccer league clubs such as English Premier League clubs use performance analysis not only to determine the rate of performance for their players but also to use the information generated from the analysis in designing a training programme aimed at addressing the team weakness and improving their strength which further makes the league more competitive and attracted many spectators (Carling et al., 2009).

Notational analysis in soccer has been in existence as far back as the 1950's, and 1960 when Reep and Benjamin in (1968) recorded and analysed data from 3213 matches between 1958 and 1968. Actions such as passing and shooting were analysed, and they concluded that 80% of goals resulted from series of three passes or less, and 50% of all goals derived from possession were acquired in the final offensive quarter of the pitch (Franks & Miller, 1986). However, over the last twenty years, a considerable amount of research has been conducted in soccer using computers (Hughes 2002; Yamanaka et al., 1997). The significant improvement in soccer performance analysis was put forward by Franks (1983) when he examined the 1982 World Cup. Likewise, Franks and Miller in (1986) utilised an options computer notational system through which six Liverpool's matches were investigated during the 1985-1986 season using a concept keyboard. Conclusions were made on patterns of play, such as a larger number of passes were undertaken while losing than winning (Hughes & Franks, 2004). Nowadays, the use of computers, videotapes, and digital cameras are widely recognised to monitor and evaluate the performance of a team or a player in a soccer match (Carling et al., 2005). In addition, Hughes and Bartlett (2002) reported that in a higher soccer league, eight cameras are set up around the stadium that allows all the twenty-two players on the field to be tracked. The data are then sent to the clubs for an in-depth analysis to be carried out on the performance of the team or an individual player.

Reliability is the extent to which measurement procedure can be relied upon to produce consistent results upon repeated application (Weiner, 2007). Sechrest (1984) defined reliability as the extent to which a test measures what it is supposed to measure in a relatively consistent manner. It can also be seen as the degree to which measures are free from blunder and consequently yield reliable results i.e. the consistency of an estimation technique (Abdullah et al., 2017b). In the event that an estimation gadget or method reliably doles out the same score to an individual or group of people with equivalent values, the instrument is viewed as dependable. Reliability include the measurement ability of an instrument to purposely evaluate what it is intended to evaluate as well as the consistency, or reproducibility, of test scores, i.e., the extent to which one can generally expect consistent deviation scores of people over testing circumstances on the same, or parallel, testing instruments (Thanasegaran, 2009).

Sports performance information has been measured and gathered via the use of various methods including different measurement devices in the area of notational analysis of sports and games (Choi et al., 2007), it is essential consequently, to investigate the reliability tests utilised for such information. Several notational analysis systems are used for notational analysis in soccer to give coaches and players information on their performance (Abdullah et al., 2015c). However, most performance analysts neglect to consider or seldom establish the reliability of their analysis tool assuming that the presence of the clear

operational definition of performance parameters ensures the reliability of the testing instrument (O'donoghue, 2007). It is vital to note that the presence of precise operational definitions neither provides excellent reliability nor does their deficiency ensure poor reliability. Establishing reliability of any performance analysis application is fundamental in order for the data to be trustworthy. Consequently, any estimation of reliability measurement thought to be adequate needs to be legitimised. Lack of reliability can bring variability into data that diminishes the possibility of discovering a critical distinction. Reliability of an evaluation tool is at any rate as imperative when performance analysis is applied in coaching and judging settings as when it is used for other scholarly research (Lames & McGary, 2007). Reliability defined, for sports performance are seen from a specific point of view that the measurement tool itself ought to yield trustworthy information (Brown & O'donoghue, 2007).

Multiple of software applications and devices are made available throughout the years to serve as the instruments for performance analysis. In a soccer match, the initial performance analysis framework was created by Reep Benjamin in 1968 when he notated a session of soccer utilising paper and pencil (Hughes, 2002). To date; several methods of the performance analysis are used to assess the performance of a player. Hand notation system is simple to use and requires less equipment's but is time-consuming and the amount of data generated by the system can involve many hours of time to process the information that can be meaningful to the coach, athlete or the sports scientist. The computerised system is more accurate and time efficient but requires a considerable learning time and very expensive. Video recording allows the match to be analysed at post events but is time-consuming and requires more expertise.

Purpose of the Research

Although many professional soccer clubs have attempted to use sophisticated performance analysis systems to analyse the performance of their players, such systems have been traditionally expensive, in addition to requiring experts to operate and maintain such systems, thus making it difficult for semi-professional and other professional clubs to afford and use them. In recent years, Mobile phone applications have grown in popularity, offering tools in the context of performance analysis that can be used to provide information on performance to both players and coaches that may help to improve their overall performances (Hook & Hughes 2001; Sermento et al., 2014). However, despite the usefulness and cost-effectiveness of these Mobile phone applications in analyzing performance, most soccer clubs do not recognize the importance of these applications in analyzing the performance of their team and players, in spite of their portability, easy in data sharing, as well as the ability to provide information to coaches on what is happening during the game in real time. In view of this, the present study aims to investigate the effectiveness of using Android-based performance analysis applications in providing information to coaches on the performance of their players. We do so using the StatWatch application—a portable and user-friendly Mobile phone application designed to monitor the performance of an athlete during a game or training. We hypothesise that the StatWatch application will reliable in providing information on players' performance to coaches in real time to understand what is going on during the game. This investigation is anticipated to be of value to semi-professional and other professional soccer clubs interested in performance analysis but are leery of cost. Specifically, the findings to be discussed herein may be of value to team owners as well as coaches.

Materials and Methods

Material's description:

StatWatch, is an Android based application for notational analysis invented by a chartered sports engineered based in Malaysia. It is an application that is compatible with a tablet or cell phones. It can code the performance of two players at a particular time with 20 accessible cells to key in the performance parameters for every player in a way that their performances can be examined simultaneously. During the selection and key in the performance parameters, the performance parameters of both the two players will show up on the screen of the tablet. A start button can then be pressed. All the required parameters to be coded would be organised on the screen, so all that is required is to tap on the parameter, and it will be consequently recorded. The data recorded can be transmitted to the coach at the required interim of the match which implies that the data can be transmitted as quick as possible to the coach as the game advances. The application has two methods for transmitting data or information to the coach. The data can be transmitted by means of Bluetooth or Email. At the end of the analysis, analysis and synopsis are provided by the application. The number of actions performed by every player, the time the action is executed, and also, the number of success or fail is recorded. Moreover, bar diagrams are displayed to project the performance of every in view of the performance parameters the players are evaluated on. Essentially, the data can further be transmitted to Microsoft Excel for additional in-depth analysis.

Figure 1: A screen shot of StatWatch application installed on a mobile phone

Participants:

 The research participants of the present study consist of one of the soccer club competing in the Malaysian Super League, eleven performance analysts and a control person (CP). The performance analyst (mean age of 24.62 and ± 3.5 years' experience of notational analysis) were recruited from a Sports Science division of University Sultan Zainal Abidin. Prior to the commencement of the study, the analysts were trained for a period of two weeks. The training included a workshop in which a real video match was shown on an LCD screen and a trip to the stadium for a real-time analysis in view to enabling the analysts' familiarise themselves with the selected performance indicators involved in the study. Similarly, a control person (C.P) with a considerable experience in performance analysis was responsible

for transmitting the data registered by the analyst to the computer after every match for further analysis.

Materials:

Eleven tablet phones installed with a StatWatch applications were used as a device for data collection.

Selection of the performance parameters:

Performance parameters were identified based on their relevant to the demand for the game and were approved by the chief coach of the club. Eleven performance analysts who were familiar with the game were trained on how to use the application and assisted in the data collection. The operational definition of each parameter was given so that, the coach, the researchers, and the performance analysts were using the terms unanimously. The players in the position of defenders, midfielders and strikers were evaluated on clearing, crossing, dribbling, heading, pursuing the loose ball, shooting, foul and through pass notated either success (s) or fail (f). Each of the performance analysts was provided with one StatWatch application to analyse the team's performances during the period under study. The players were observed during their games, and their performances were recorded based on the performance indicators already selected. The performance of the team for four weeks was analysed.

Reliability Testing Procedures:

The eleven performance analysts were required to notate the performance of the team for four competitive matches consecutively via the application and all the recorded actions were collected to test the agreement between all the analysts based on the total actions registered as well as the consistency of the application to provide information on the team performances within the four matches played. This is deemed necessary in order to test whether the application has the ability to provide information on the players' performance at different times and in a relatively consistent manner.

Setting: The study took place in Malaysia during 2014-2015 Malaysian Super League Season.

Data Analysis

Absolute reliability values (mean, standard deviation, standardised typical error and the intra-class correlation coefficient) of numerous individual actions of the players were calculated utilising the spreadsheet established by Hopkins (2000). All team performances registered by the eleven performance analysts were analysed. Scatter plots were drawn to illustrate the summary for the total actions performed by the players in the four matches played as recorded by the performance analysts. The data for the total of four games were analysed and evaluated using spreadsheet established by Hopkins (2000) and XLSTART add in version 2014 for Windows.

Results

Table 1:

Descriptive statistics.

Observers	M	SD	N(Indicators)
Analyst1	18.0	29.8	20.0
Analyst2	18.7	22.0	20.0
Analyst3	21.6	30.1	20.0
Analyst4	22.3	25.7	20.0

Analyst5	23.0	30.9	20.0
Analyst6	23.0	45.4	20.0
Analyst7	23.8	38.8	20.0
Analyst8	40.6	53.3	20.0
Analyts9	18.7	22.0	20.0
Analyst10	21.6	30.1	20.0
Analyst11	22.3	25.7	20.0

Table 1 shows the descriptive statistics of the variables involved in the study. The number of the analysts, the mean, standard deviation and the number of performance indicators selected are displayed.

Table 2:
Reliability statistics.

A. M. for ICC	95% CI Lower Bound	95% CI Upper Bound	F	df	Sig.	Cronbach's Alpha	No of observers
0.98	0.97	0.99	67.57	19.00	.000*	0.99	11.00

Abbreviations: A.M for ICC= Average measure for the intra-class coefficient, CI= Confidence interval.

Table 2 demonstrates the inferential statistics for the reliability test. The intra-class average measure reveals 0.98 which explains that 98% of the variance and the means of the analysts in their overall analysis is real $p < 0.001$. Similarly, a Cronbach's Alpha coefficient indicates .099 which suggests an excellent consistency in their general observation.

Table 3:
Reliability for the players' actions in the four matches played.

Performance Indicators	Mean	Standard Deviation	Standardized typical error	Intra-Class Correlation (ICC)
Crosses Successful	39.27	14.46	0.36	0.87
Crosses Fail	24.36	21.14	0.37	0.93
Clearing Successful	23.36	21.54	0.49	0.83
Clearing Fail	24.00	21.05	0.35	0.90
Dribbles Successful	33.55	17.54	0.29	0.91
Dribbles Fail	26.18	19.89	0.6	0.94
Heading Successful	39.00	14.59	0.4	0.83
Heading Fail	34.45	17.39	0.24	0.94
Chasing Loose Ball Successful	26.18	20.27	0.11	0.92
Chasing Loose Ball Fail	34.82	24.41	0.36	0.75
Open-space Successful	32.91	17.72	0.34	0.80
Open-space Fail	23.64	21.84	0.59	0.87
Passes Successful	109.09	77.19	0.27	0.40
Passes Fail	40.09	16.30	0.91	0.80
Shots Successful	20.73	23.60	0.12	0.87
Shots Fail	25.36	20.09	0.06	0.91

Fouls Successful	23.45	21.60	0.51	0.91
Fouls Fail	25.09	20.41	0.15	0.89
Through Pass Successful	23.27	21.76	0.56	0.87
Through Pass Fail	21.82	22.74	0.86	0.86

Table 3 reveals the absolute reliability of the performance indicators recorded by the analysts for the player's performances in the matches played. It can be observed that the standardised typical errors range from 0.04 to 0.86 while the intra-class correlation coefficients vary from 0.40 to 0.94 revealing an excellent and acceptable reliability for the overall analysis.

Table 4:
correlations of the agreements among the performance analysts.

Observers	1	2	3	4	5	6	7	8	9	10	11
Analyst 1	1										
Analyst 2	0.97	1									
Analyst 3	0.97	0.96	1								
Analyst 4	0.90	0.92	0.91	1							
Analyst 5	0.92	0.91	0.95	0.91	1						
Analyst 6	0.99	0.96	0.97	0.89	0.91	1					
Analyst 7	0.95	0.93	0.93	0.93	0.89	0.96	1				
Analyst 8	0.96	0.97	0.93	0.95	0.91	0.95	0.94	1			
Analyst 9	0.97	1.00	0.96	0.92	0.91	0.96	0.93	0.97	1		
Analyst 10	0.97	0.96	1.00	0.91	0.95	0.97	0.93	0.93	0.96	1	
Analyst 11	0.90	0.92	0.91	1.00	0.91	0.89	0.93	0.95	0.92	0.91	1

Table 4 indicates the relationships of the agreement among the performance analysts. It can be seen from the table that there are comparatively perfect positive agreement between the performance analysts. The range of the agreement across all the analysts varies from 0.89 to 1.0 which further reveals a strong agreement among them.
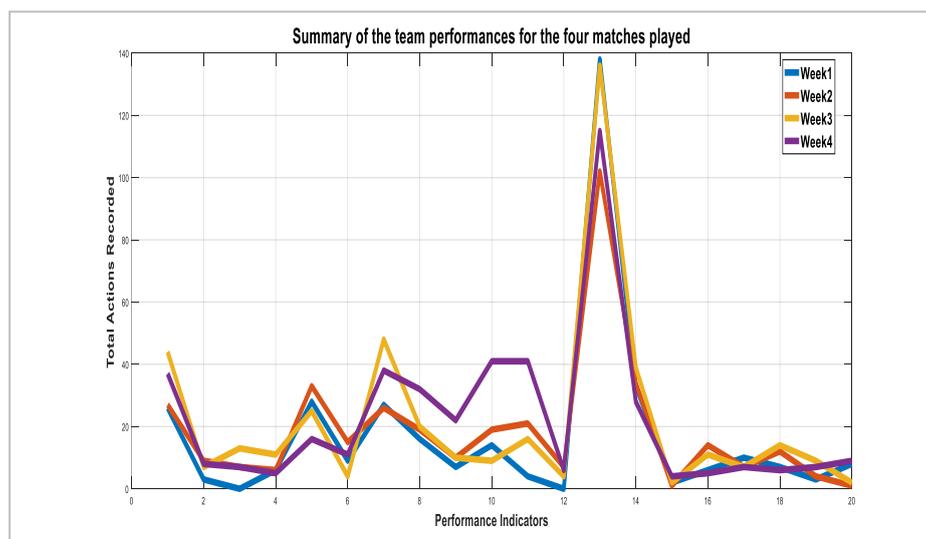

Figure 2: Summary of the total actions performed by the team in the four matches played

Figure 2 displays the total actions performed by the players in the four matches played as recorded by the performance analysts. From the figure, it can be observed that the overall

performances of the team in between the weeks have not followed a similar pattern. This revealed that the team performed differently in between the weeks which could be resulted from the change of the strategies as well as the particular demand and difficulties of each game.

Discussion

The current study reveals that the method of notational analysis using the instrument of observation *StatWatch* has a considerable inter-observer reliability. The results gathered by the tool of the analysis can be valid for subsequent research studies (Table 2). Advances in Mobile phone application technology in recent years have led to a wide variety of systems that football managers are now using to make and implement decisions. These systems have been developed from scratch for specific purposes and differ significantly from one another in terms of their ability to provide information to their end users. It is, therefore, pertinent and necessary for every application to be tested in order to ascertain whether it has the capability to evaluate what it is initially designed to evaluate (Gay, 1987).

Conversely, Table 3 has revealed the reliability for the players' actions in the four matches played which has further stressed the strength of the application in providing information on performance. This finding is supported by Berk, (1979) who pointed out that reliability is arguably the most important criteria for testing the quality of any evaluating instrument. The term reliability here denotes to whether or not the test measures repeatedly what it asserts to evaluate. However, referring to a test with a noticeable reliability, the items will be closely connected to the test's proposed application. In several certification and licensure tests, the items will retrieve consistent results related to a specific intention or purpose to what it is designed. If a test has poor reliability, however, then it does not measure consistently the related constructs and capabilities it should. Therefore, once this is the case, there is no justification for using the test results for their intended purpose.

There is also sufficient evidence to claim that StatWatch application is able to make reliable measurements of the performance of the team when we reflect the correlations of the agreements among the performance analysts (Table 4). The application can measure the performance of the team in a consistent manner. Ertan et al., (2005) reported that reliability is among of the most important elements of device's quality. It has to do with the consistency or repeatability of an instrument's performance on the test. For instance, if we were to administer a test with high reliability to an application on two different instances, it could be very likely to reach the same conclusions about the instrument's performance both times. A test with less reliability, on the other hand, might result in very different scores for the examinee across the two test administrations. If a test produces inconsistent scores, it may be unethical to take any substantive actions on the basis of the test (Carmines & Zeller, 1979). However, the finding from this study has indicated that the overall performances of the players in between the weeks have not followed a similar pattern. This revealed that the players performed differently in between the weeks which could result from the change of the strategies as well as the specific demand and difficulties of each game (see Figure 2).

Conclusion

 Among essential features for any performance, analysis devices are the ability to generate information, analyse the information in a way that is comprehensible to both the coach and the player as well as the user-friendliness of the device. This study revealed that notational analysis in soccer using Android based application is effective in providing the coach with

information on the performance of the team. The nature of the application such as portability, user friendliness and faster in sharing information made it easier to notate players performance with the purpose of accelerating their development and achieving success. It is our hope that professional and semi-professional soccer clubs will consider Android based application in performance analysis of their teams due to its effectiveness in the notational analysis.

Limitations and future direction
StatWatch application does not have the capacity in the motion analysis of players; a cost effective application should be developed to monitor the movements of the players during the performance. Other studies might be required to compare the validity and user-friendliness of StatWatch application with other computerised based notational analysis systems.

Acknowledgement

Corresponding Author
Mohammad Razali Abdullah,
Faculty of Applied Social Science, Universiti Sultan Zainal Abidin, 21300 Gong Badak Campus, Kuala Terengganu, Terengganu, Malaysia. razali896@yahoo.com

References
Abdullah, M. R., Haque, M., Musa, R. M., Kosni, N. A., Maliki, A. B. H. M., & Suppiah, P. K. (2017b). Establishing Reliability in Notational Analysis of Soccer of Malaysia. *Int. J Pharm Sci Rev Res, 43*(2), 164-168

Abdullah, M. R., Mu'azu R. M., Suppiah, P., Akram, M. B., Tomny, W.K.K., and Mubin C, Z. (2015c). Comparative Analysis on the User-Friendliness between Computer and Tablet Application in the Performance Analysis of Soccer. *International Journal of Science and Technology, 4*(1) 368-372.

Abdullah, M. R., Musa, R. M., Maliki, A. B. H. M. B., Kosni, N. A., & Suppiah, P. K. (2016a). Role of psychological factors on the performance of elite soccer players. *Journal of Physical Education and Sport*, *16*(1), 170.

Berk, R. A. (1979). Generalizability of behavioral observations: a clarification of interobserver agreement and interobserver reliability. *Am J Ment Defic*, *83*(1), 460-472.

Brown, E. and O'Donoghue, P. (2007). Relating reliability to analytical goals in performance analysis. *International Journal of Performance Analysis in Sport, 7*(1), 28-34.

Carling, C., Relly, T. and Williams, A. (2009). *Performance assessment for field sports*. London, Routledge.

Carling, C., Williams, A. and Relly, T. (2005). *The handbook of soccer match analysis*. London, Routledge.

Carmines, E. G., and Zeller, R. A. (1979). *Reliability and validity assessment,* (Vol. 17). Sage publications.

Choi, H., O'Donoghue, P., and Hughes, M. (2007). An investigation of inter-operator reliability tests for real-time analysis system. *International Journal of Performance Analysis in Sport, 7*(1) 49-61.

Ertan, H., Kentel, B. B., Tümer, S. T., and Korkusuz, F. (2005). Reliability and validity testing of an archery chronometer. *J Sports Sci Med, 4*(2), 95-104.

Franks, I. M., and Miller, G. (1986). Eyewitness testimony in sport. *Journal of sports behavior*.

Franks, I.M. (1983). Analysis of association football soccer. *J Sports Sci, 9(*1), 35-43.

Gay, L. R. (1987). *Educational Research: Competencies for analysis and application*. Columbus, Merrill Pub. Co. pp.180-185.

Hook, C. and Hughes, M.D. (2001). *Patterns of play leading to shots in Euro 2000*. In Pass.com. Ed, CPA (Center for Performance Analysis). Cardiff: UWIC. pp. 295-302.

Hopkins, W. G. (2000). *Reliability from consecutive pairs of trials (Excel spreadsheet).* A new view of statistics. Sportsci.org: Internet Society for Sport Science. Sportsci.org/resource/stats/xrely.xls; [Accessed 6-11- 2016].

Hughes, M, and Franks I.M. (2004). *Notational analysis of sport: Systems for better coaching and performance in sport*. 2nd Edition, London, Routledge.

Hughes, M. (2002). *The application of notational analysis to racket sports. Science and racket sports* II. London, Routledge.

pp. 211.

Hughes, M. D., and Bartlett, R. M. (2002). The use of performance indicators in performance analysis. *J Sports Sci, 20*(1), 739-754.

Lames, M, and McGarry, T. (2007). On the search for reliable performance indicators in game sports. *International Journal of Performance Analysis in Sport., 7*(1), 62-79.

O'Donoghue P. (2010). *Research method for sports performance analysis*. London. Routledge.

O'Donoghue, P. (2007). Reliability issues in performance analysis. *International Journal of Performance Analysis in Sport, 7*(2), 35-48.

Sarmento, H., Marcelino, R., Anguera, M. T., Campaniço, J., Matos, N., and Leitão, J. C. (2014). Match analysis in football: a systematic review. *J Sports Sci., 32*(20), 969-974.

Sechrest, L. (1984). *Reliability and validity. In* Bellack AS, Hersen M, Ed. Research methods in clinical psychology. USA. pp.24-54.

Thanasegaran, G. (2009). *Reliability and validity issues in research*. Integration & Dissemination, 4, pp 35-40.

Weiner J. (2007). *Measurement: Reliability and Validity Measures*. [ONLINE] Available at: http://ocw.jhsph.edu/courses/hsre/pdfs/hsre_lect7_weiner.pdf. [Accessed 23-3-2015].

Yamanaka, K. U., Liang, D. Y., and Hughes, M. (1997). *An analysis of the playing patterns of the Japan national team in the 1994 World Cup qualifying match for Asia*. Science and football III. pp. 221-228.