Effects of Peer-Assessed Feedback, Goal Setting and a Group Contingency on Performance and Learning by Academy Youth Soccer Players

Josh E. Holt, Gary Kinchin, Gill Clarke

University of Southampton

Paper presented at the AIESEP 2011 International Conference,

UL, Limerick, Ireland, 25 June 2011

Effects of Peer-Assessed Feedback, Goal Setting and a Group Contingency on Performance and Learning by Academy Youth Soccer Players

Coaches responsible for talent development in team sports need to maximise practice and learning of essential skills of the game and accurately and continuously assess the performance and potential of each player. The relative age effect highlights the erroneous process of initial and on-going player assessment, based largely on the subjective opinion of game performance by coaches (Jiménez & Pain, 2008; Williams & Riley, 2000). Two strategies for a more objective assessment of performance are skills tests and statistics from competition such as the Team Sport Assessment Procedure (Gréhaigne, Godbout & Bouthier, 1997). However, an objective assessment of each player’s progress based on the repeated measurement of essential skills during practice is worthwhile.

A recent study showed that objective feedback from parent observer-assessors with goal setting and a group contingency improved effort, performance and learning by young soccer players during individual practice (Holt, Kinchin & Clarke, in press). Group contingencies are effective strategies to maximise the behaviour of a team and involve the award of a common, positive consequence dependent on individual goal attainment. In group practices, players could be used to assess peer performance. Peer assisted learning strategies are effective tools for physical education teachers to enhance feedback and learning (Ward & Ah-Lee, 2005) but research on the use of peers by coaches is sparse. The purpose of this study was to measure the effect of peer-assessed feedback, goal setting and a group contingency on performance and learning in a group technical practice. A second purpose was to measure each player’s passing, first touch and an awareness response and to measure the reliability of peer-assessment as an immediate source of objective feedback.

**Method**

**Participants and setting**

Participants were five, 10 to 12 year old boys who played for an English professional soccer club academy. I was the coach for all sessions, am a UEFA ‘A’ Licence holder and an experienced PE teacher. Two practices occurred during weekly, indoor coaching sessions for a total of 22 practices during the second half of a soccer season.

**Dependent variables and data collection**

Players performed a technical practice called the ‘passing square’ to help them become highly skilled at passing and receiving the ball with both feet. The practice shown in Fig. 1 involved six or more players passing two soccer balls at pace and moving around the outside of a 10 yard square. Each player completed 20 passes using his right foot followed by his left foot in the opposite direction. Prior to the study the coach and players defined eight skills or behaviours required in the practice and how to perform each one successfully. A detailed task analysis was written down on flip chart paper and displayed on the wall nearby. The three most important skills were selected as dependent variables for the experiment; an awareness response, passing and first touch.

‘Awareness’ before receiving the ball was a head turn with a focused look to check the position and movement of the support player after the ball had been received by the player making the pass. Four criteria defined a successful pass; the use of the correct foot, a clean foot to ball contact and the accuracy and firm weight were completely defined. A successful first touch occurred when the player used the correct foot to position the ball so he could pass within three strides with his next touch using the inside of the foot. Unsuccessful responses were also defined as were all situations with no opportunity to respond. Baseline data were collected from video recording of practice and intervention data were collected from recording and live by a player and the coach. Observers assessed every response made at one corner of the square, recording 1 for successful or correct, 0 for incorrect and X if the player had no opportunity to respond.

**Procedure, intervention and research design**

The coach and players quickly reviewed the main task requirements before the start of every practice. The first practice of each session was after the warm up, the second occurred towards the end of the session and both were immediately followed by a 10-min 4v4 game. Encouragement and praise for effort was given periodically to the group to help sustain a high practice tempo.

**Intervention (B).**After the first 7 baseline practices, the coach met with the players and parents to give each player a Performance and Learning Chart (PLC) containing his percent correct scores for awareness. The intervention consisted of peer-assessed feedback with charting, goal setting and a group contingency. Players were taught to observe, peer-assess and record performance scores using video of baseline practices. Training occurred until the majority of players had an agreement with the primary data greater than 80%. Players self-set personal goals and following agreement with the coach recorded a horizontal goal line on their chart. The group contingency involved the award of 10-min of extra game time if half the group achieved their goals. Practice and coaching were as baseline except for the intervention procedures. Each player was given objective feedback as a percent correct score to plot on his chart immediately after the practice and when the majority of scores indicated a stable trend the intervention was applied to passing followed by first touch. A single subject, multiple baseline experiment was used to assess the effects of the intervention on performance of the three skills.

**Interobserver agreement and social validity**

Agreements were calculated on 36% of primary video data across all experimental phases and conditions and 100% of live intervention data to assess the reliability of performance measures. For primary data the scores counted by an independent observer, who was an experienced physical education teacher and coach were compared to the scores counted by the primary investigator. For intervention data the peer-assessed scores were compared with the primary data. Agreements were calculated using the trial-by-trial method and dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100% (Cooper, Heron & Heward, 2007). Two social validity assessments, one during and another after the experiment evaluated the player’s satisfaction with the goals, procedures and their learning outcomes.

**Results**

Results for two players shall be presented with an indicative description for Charlie. Figure 2 shows Charlie’s practice data with dots and squares representing right and left foot scores respectively. During baseline or condition A his awareness was very low with a slight improvement when using his right foot. At the start of the intervention or condition B his awareness immediately became complete and remained at 100% correct throughout. Peer-assessed scores that do not correspond exactly with primary data are shown in red to give a visual, reliability comparison. Charlie’s passing during baseline was initially stable and moderate with his right foot, which then improved noticeably before dropping in practice 12. His left foot passing showed a steady improvement but was variable for the last two baseline practices. During the intervention his passing with both feet improved and remained stable above his goal of 70% correct for all except one practice. His first touch during baseline was initially highly variable with his right foot followed by a steady improvement to 80% correct in practice 18. His left foot showed considerable improvement over the first six practices after which his performance dropped and was somewhat variable. During the intervention his first touch with both feet remained stable and above his goal of 80% correct for all except one practice, with his left foot particularly high ranging between 89 and 100% correct.

The reversal to baseline for awareness and passing provided a measure of maintenance after the intervention was removed. Charlie maintained awareness initially above or close to his 80% goal; it then dropped towards the end of the study but remained noticeably higher than baseline performance. His passing remained above his 70% goal with both feet. Figure 3 shows James’ data which indicate variable performance during baseline and improved and more stable performance during the intervention and represents general trends for the group.

**Interobserver agreement**

Table 1 shows agreement scores indicating the reliability of primary and peer-assessed data. The overall mean agreements were 83.4% for the primary data and 82.2% for the peer-assessed data. The 58% agreement for awareness was from the first intervention practice when the peer-assessor failed to account for one criterion. The next lowest agreement was 77%.

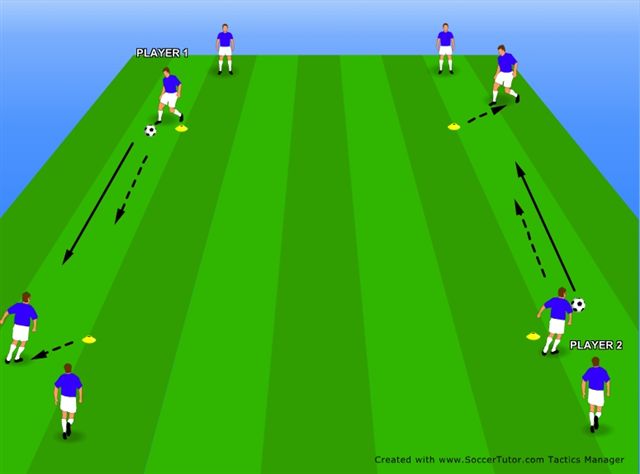
**Discussion**

This experiment has demonstrated positive intervention effects on the quality of performance and learning of three essential skills. When players were; (1) required to define and write down exactly what they had to do, (2) asked to set challenging and measurable goals, (3) made formally accountable for their performance with objective peer-assessment and self-charting and (4) rewarded as a group for individual goal attainment they performed consistently better in practice which improved learning. This supports the primary finding from ecological physical education research that accountability drives the instructional task system (Siedentop, 2002). As one player commented; “I was motivated by knowing my scores were going to be recorded. Also what made me do better was that people were watching me”.

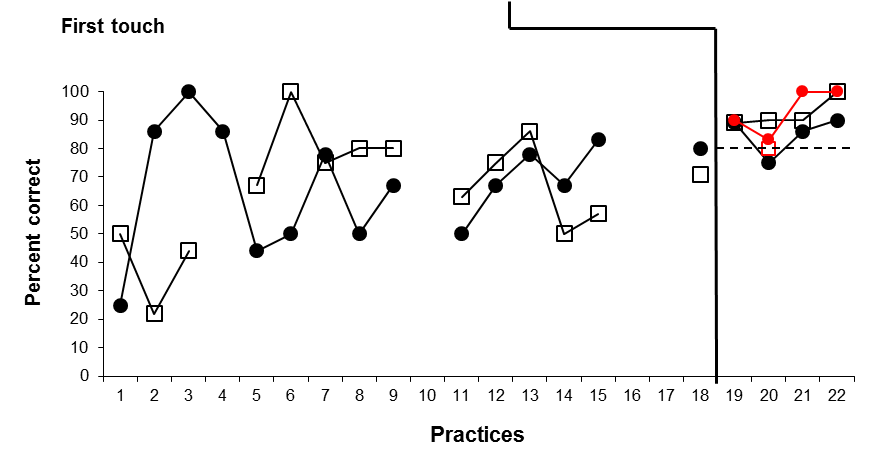
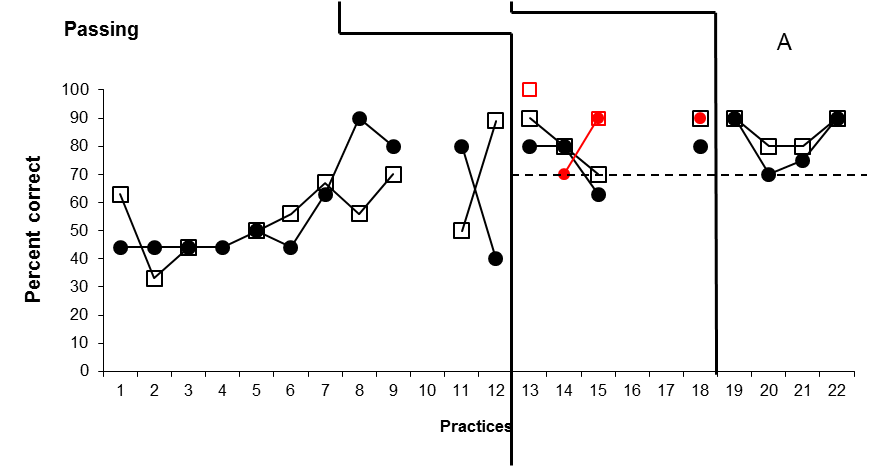
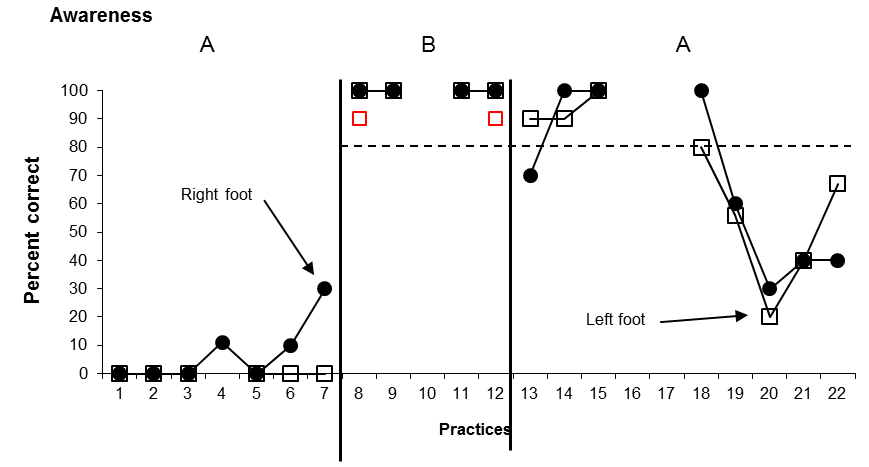
The effectiveness of any instructional strategy is dependent on learning outcomes and the learners’ satisfaction with the procedures and outcomes. The players reported an increased confidence in these skills in competitive situations but Fig. 4 shows their enjoyment of the practice decreased sharply during baseline, which reflects previous research findings that deliberate practice is not inherently enjoyable (Holt, Kinchin & Clarke, 2008/ in press; Ericsson, Krampe and Tesch-Romer, 1993). An increased approval of the practice is seen as the intervention was applied to the second and third dependent variables, possibly due to the combination of receiving scores and improved competence. This study shows how a coach can encourage deliberate practice with the reward of bonus game time. The players generally enjoyed acting as peer-assessors, although the effect of this on their subsequent performance and learning is not clear and warrants future research.

The second purpose was to examine the reliability of peer-assessment for objective feedback and performance data. The mean and range agreement scores from the most conservative trial-by-trial calculation demonstrates the ability of these young players to reliably assess and record peer performance of three essential skills of the game. With a total of nine players acting as peer-assessor, these agreements provide firm confidence that any variability in the data reflect real variations in practice performance (Cooper, Heron & Heward, 2007). Coaching implications of the peer assessment process include the time and preparation required to train the players to observe and score accurately. Future research needs to examine the ability and willingness of other coaches to use similar methods and to evaluate the congruent validity of data to discriminate the most able players. In regular coaching the reliability and honesty of peer-assessed data could be checked by the coach once peer scores reach a criterion level. Limitations of this research include the use of two practices per session and the motivational issues caused by extended baselines. To sustain motivation coaches could set a maximum number of practices to achieve established performance standards or benchmarks. The lack of treatment integrity assessment was also a limitation.

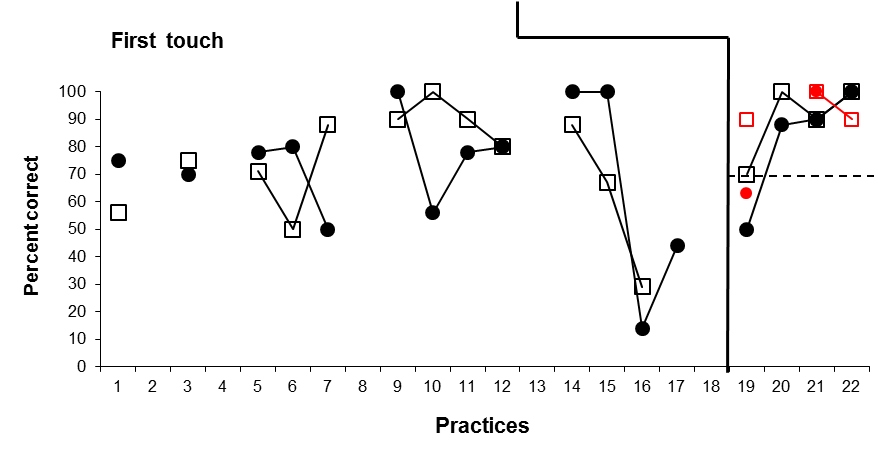
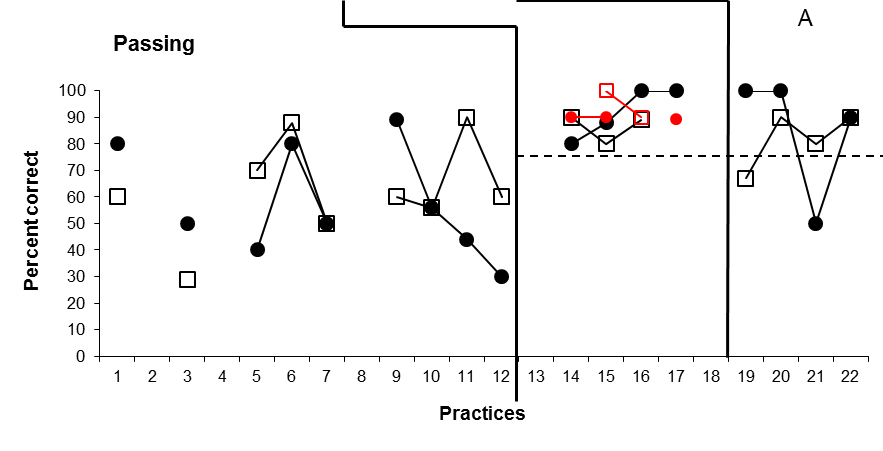
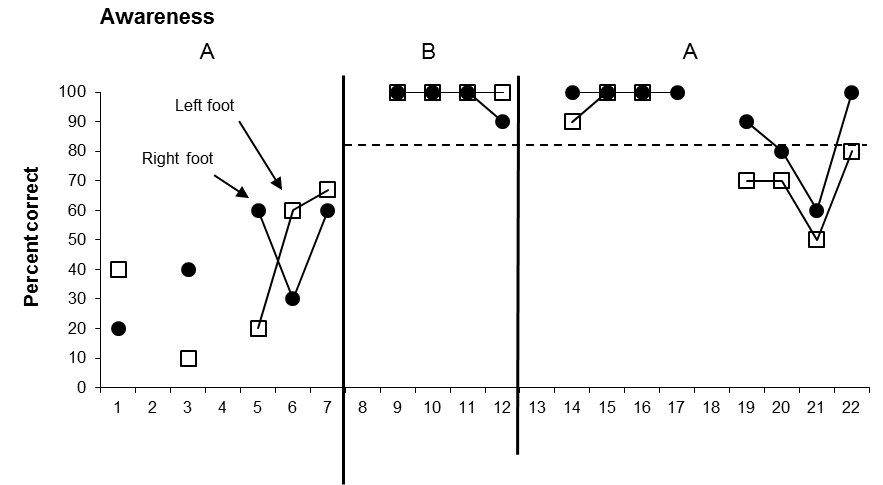
In summary, this is the first study to demonstrate methods to measure and chart practice improvements in essential skills during talent development coaching. Self-set goals and charting peer-assessed data with a group contingency successfully improved the awareness, first touch and passing by all players. The findings suggest that this accountability, feedback and reinforcement procedure provides a promising method to move children forwards, to know where they have come from, where they are in relation to where they need to get to and how to get there. Implications for future research with more complex learning tasks and for coach education and practice include the potential for repeatable and measurable systems to track performance and learning to help develop and determine the most talented players.



*Figure 1.* Diagram of the ‘passing square’ drill practice (solid arrows indicate a pass, dashed arrows indicate player movement).



*Figure 2.* Practice data for Charlie. Percentage of correct awareness, passes and first touches with his right foot and left foot



*Figure 3.* Practice data for James. Percentage of correct awareness, passes and first touches with his right foot and left foot



*Figure 4.* Mean player enjoyment scores for the passing square practice (1 = ‘boring/ dull’; 2 = ‘not much fun’, 3 = ‘OK’, 4 = ‘some fun’, 5 = ‘great fun’)

*Table 1.* Mean and range interobserver agreements for all three dependent variables

|  |  |  |
| --- | --- | --- |
| **Dependent variable** | **Primary-video data (baseline & maintenance)** | **Intervention** |
| **Awareness (A)** | 87.8% (range, 75-98%) | 84.8% (range, 58 to 100%) |
| **Passing (P)** | 79.8% (range, 73-90%) | 80.7% (range, 76 to 87%) |
| **First Touch (FT)** | 82.6% (range, 78-87%) | 80.3% (range, 73 to 90%) |
| **Overall** | 83.4% | 82.2% |