Inter-Individual Variability of Teachers’ Feedback. Study in Simulated Teaching Conditions

Marc Cloes, Alain Denève, & Maurice Piéron
University of Liège, Belgium

In physical education teaching and in classroom teaching, most teachers are convinced of the role of feedback in the pedagogical context (Bloom, 1979; Rosenshine, 1980). Feedback has been identified as an important behavioural descriptor of the teacher’s enthusiasm (Caruso, 1980; Cloes & Piéron, 1989; Rolider, 1979). Its positive influence on pupils’ learning has been evidenced several times (Cerreiro da Costa & Piéron, 1992; De Knop, 1986; Phillips & Carlisle, 1983). However, some authors consider that its influence might be overestimated (Lee, Keh & Magill, 1993; Magill, 1994; Silverman, 1994), but significant differences were observed when experts’ and novices’ feedback were compared (Piéron & R. Delmelle, 1983; Piéron & V. Delmelle, 1983). Despite such critiques, feedback remains a key interest in effective teaching and should be further investigated in natural and quasi experimental conditions.

For some considerable time, improving the delivery of feedback was an important objective of programmes of behaviour modification in pre- and in-service teachers (O’Sullivan & Burrough, 1989; Piéron & Wauquier, 1984; Siedentop, 1981). Several descriptive studies showed that quantitative (ratio, rate and frequency), structural (intent, form, direction), and qualitative (specific referent, adequacy) aspects of feedback were characterized by large interindividual variability (Fishman & Tobey, 1978; Piéron, 1982). The mechanisms of feedback provide insight in understanding its interindividual variability. Several authors have proposed models of feedback (Armstrong & Imwold, 1982; Hoffman, 1983; Pinheiro & Simon, 1992).

Following the concept developed by Hoffman (1983), they emphasized the existence of two important steps in the feedback process:

(i) the diagnostic: analysis of the performance and identification of the nature, extent and cause of the discrepancy between the learner’s actual and desired responses;

(ii) the prescriptive: selection and application of a remedy to the discrepancy.

By definition, feedback is 'a teaching behaviour dependent upon the major response of one or more students and intended to provide information related to the acquisition or performance of a motor skill' (Fishman & Anderson, 1971, p. 11). Obviously, the learner’s performance characteristics are prevailing factors influencing the intra- and inter-individual variability of the feedback. Nevertheless, other variables also seem to be involved in the process of feedback. Individual variables like the teacher’s experience and knowledge of the task are influential. This has been identified through comparison of feedback given by specialists and non-specialists (Faucette & Paterson, 1990). Moreover, planned decisions can also modulate feedback, as has been shown by Cloes, Zabus & Piéron (1991).

Thus, feedback process models should consider the influence of this variable (refer figure 1).

To shed light on teachers’ decision making and to figure out the role of factors like teachers’ diagnosis, proactive decisions or personal experience, it is necessary to set up quasi-experimental designs in which teachers react to identical motor performances. Simulated teaching situations
provide suitable conditions. Teachers have to react to pupils’ performances shown on video-tape as if they were in a natural situation (Cloes, Périon, Colomberotto, Baret & Brouwers, 1988; Katz & Cain, 1985).

The purpose of the study reported in this article was to identify the main sources of inter-individual variability of feedback emitted by teachers reacting to identical pupils’ motor performances.

Method
A set of video recording sequences was developed to prepare a standardized simulated teaching session. The sequence used was a series of volleyball spikes performed by intermediate skill level female
players. The session was followed by one individual at a time. It comprised a training period which enabled subjects to familiarize themselves with the simulated teaching technique. During that period, sequences were presented until the subject had reached more than 90% of direct interventions, i.e. interventions issued to the pupils appearing on the screen. It is noteworthy that only one out of all subjects mentioned a feeling of discomfort in the simulated teaching setting. After the training period, the subject watched the sequence aimed at the analysis of feedback variability. It was a 15-minute video recording showing a sequence of 110 spikes.

Feedback was recorded for subsequent analysis through a multi-dimensional observation system adapted from an instrument developed and used at the University of Liège (Piéron & R. Delmelle, 1983). Three dimensions were considered: intent, referent and adequacy (refer table 1). The rate of feedback was also calculated.

**Table 1.**

The multi-dimensional observation system

- **Intent: Identification of the feedback purpose**
  - Global approval feedback: intended to provide a favourable appraisal of the whole performance
  - Specific approval feedback: intended to provide a favourable appraisal of a part of the performance
  - Global disapproval feedback: intended to provide an unfavourable appraisal of the whole performance
  - Specific disapproval feedback: intended to provide an unfavourable appraisal of a part of the performance
  - Descriptive feedback: intended to describe the performance
  - Prescriptive feedback: intended to provide instructions for collecting subsequent performances
  - Interrogative feedback: intended to engage the learner in a reflection on his/her performance
  - Affective feedback: intended to provide an attitudinal or motivational set towards the performance

- **Referent: Identification of the feedback content**
  - Non-specific feedback: focused on the result of the performance
  - Feedback focused on the ‘set-up’: distinction between five parts of the skill
  - Feedback focused on the ‘spike’: distinction between six parts of the skill, comprising 14 specific criteria

**Adequacy: Identification of the consistency between the feedback message and pupils’ performance**

- Appropriate: high level of consistency
- Inappropriate: very low level of consistency
- Undetermined: consistency level impossible to find out

Twelve subjects took part in the study. They were distributed in four groups of three according to teaching/coaching experience and specialization level in volleyball (refer table 2).

**Table 2. Characteristics of the four groups**

| Group A: Specialist student teachers (S1, S2, S3) |
| At least five years of volleyball competition at an intermediate or high level |
| Any teaching experience |
| At most two years of volleyball coaching (youth teams) |

| Group B: Specialist teachers (S4, S5, S6) |
| At least five years of volleyball competition at an intermediate or high level |
| At least five years of teaching experience |
| At least five years of volleyball coaching (youth, intermediate or high level teams) |

| Group C: Non-specialist student teachers (S7, S8, S9) |
| Any experience of volleyball competition |
| Any teaching experience |
| At most two years of coaching (youth teams) in a sport other than volleyball |

| Group D: Non specialist teachers (S10, S11, S12) |
| Any experience of volleyball competition |
| At least five years of teaching experience |
| At least five years of coaching (youth, intermediate or high level teams) in a sport other than volleyball |

In this report, subjects characterized by the same volleyball experience (specialization level), are termed ‘specialists’ or ‘non-specialists’. When considering the pedagogical experience, the term ‘teachers’ designates all subjects having substantial teaching/coaching experience whereas ‘student teachers’ is used to name subjects with minimal experience; specialist teachers who combine teaching experience, coaching experience, and participa-
tion in volleyball competition are termed "experts".

Intra-observer reliability exceeded 0.85 of agreement (Bellack). Data were processed by a non-parametric statistical test for comparison between groups (Mann Whitney U Test). A computerized statistical package developed by the Laboratory of Experimental Pedagogy at the University of Liège (ADDAD) was used for cluster analysis.

**Results**

Data were processed in two steps: 1) to compare feedback characteristics in the four groups; 2) to identify classes of subjects providing similar types of feedback.

**Feedback characteristics in experimental groups**

1. **Rate of feedback**

Subjects reacted very frequently to pupils' motor performances. The rate averaged 12 feedback episodes per minute, i.e. a frequency of one feedback every five seconds. The high rate of feedback was similar to observations in comparable situations when teachers interacted with small groups, during short periods and when the purpose of the session was to provide feedback to learners (Cloes, Zabus & Piéron, 1991). High feedback rates were also observed in an earlier simulated teaching experiment (Cloes et al., 1988). Individual's feedback frequency ranged from one feedback every three seconds to one every 10 seconds.

The specialization level did not seem to influence the quantitative aspect of feedback. This finding corresponds with previous results. The relationships between the specialization level and the feedback rate showed low stability (Armstrong, 1986).

The teaching/coaching experience could be a more powerful (or meaningful) variable than specialization in interpreting the variability of the feedback rate. Teachers tended to issue more feedback than student teachers: 14 Vs 10/min (U = 11; p = .155). This confirms results from previous studies in natural situations (Fishman & Tobey, 1978; Piéron & R. Delmelle, 1983; Piéron & V. Delmelle, 1983) or in simulated teaching settings (Cloes et al., 1988). Moreover, the three subjects with the most important teaching/coaching experience (S5, S6, and S12) issued the highest rates of feedback. Two of them were specialists in volleyball, the third in gymnastics. This result underlined the possible influence of teaching/coaching experience on the feedback rate.

2. **Intent of feedback**

The profile of intent categories was similar to that observed in other natural situations (Piéron, 1982). It was characterized by the highest proportion of prescriptive feedback (53.4%) and nearly similar proportions of approval/disapproval (18.2%) and descriptive (25.1%) feedback. Interrogative and affective feedback averaged 3.3%. Descriptive feedback was less frequent than observations in a previous simulated teaching experiment (Cloes et al., 1988). This is probably due to the improvement of the training procedure used in this research in the simulated teaching session.

The variability of the distribution of intent categories was fairly large. It seemed not to be influenced by teaching/coaching experience. On the other hand, specialization could induce more variability. It was illustrated by a larger use of specific approval/disapproval feedback by the specialists than by the non-specialists: 48 Vs 26.7% (U = 6, p = .032).

Non-specialists seemed to lack the necessary knowledge of the task. Thus, they could be limited to issue frequent stereotyped feedback. This finding was similar to a default option used by less effective teachers (Yerg & Twardy, 1982).

3. **Referent of feedback**

Describing the referent of feedback was the first step to appraise the quality of the feedback. Analysis of feedback referent facilitated the identification of the element that subjects considered as relevant to improve the quality of pupils' performances. It also provided interesting information about the process of feedback.

The proportion of feedback involving content information was found. Specific feedback averaged 86.9%. In this category, the ranking of subjects was as follows: (i) in-service teachers (90%);
(iii) specialist student teachers (87%); (iii) non-specialist teachers (85.5%); (iv) non-specialist student teachers (83.9%). Subjects with the highest level of specialization tended to provide more specific feedback than non-specialists: 88.8 Vs 84.3% (U = 10; p = .12).

These results confirm previous observations on the specificity of approval/disapproval feedback. Combining a high level of specialization and teaching/coaching experience would help physical educators to become more accurate in task analysis and in selecting the information enabling learners to improve their performances. This confirmed findings from Harari & Siedentop (1990), Girardin & Hanson (1967), and Bard, Fleury, Carrière & Hallé (1980).

All subjects were more often concerned by the 'spike' than by other categories. No particular tendency was observed between groups. However, although not statistically significant, teachers balanced their feedback more between 'spike' and 'set' than did student teachers: 59.4 and 33.7% Vs 65.4 and 27.8% (U = 15; p = .350).

When dealing with individual cases, subjects' role in the game and feedback referent characteristics happened to be related. For example, a specialist student teacher playing as setter issued more feedback focused on the set up while a specialist teacher playing as hitter was concerned mostly by the spike.

The specific interest of specialists could interfere with the development of a balanced frame of reference. Subjects 1 and 4 were the most concerned by the game as players when they participated in the experiment. One can hypothesize that when educators think more like players than like teachers/coaches, their feedback process is influenced by different pre-active decisions. This finding underscores the complexity of the mechanism of the feedback.

At first sight, teaching/coaching experience would lead to a balanced distribution of feedback between the different parts of the movement. This was particularly underscored in the feedback focused on the spike (refer table 3). In comparison with student teachers, teachers focused 24.4 Vs 16.5% on the hit preparation (U = 5; p = .021) while student teachers were more concerned than teachers by the 'hit': 34.3 Vs 23.6% (U = 6; p = .032).

Table 3. Ranking of the main categories of spike referent feedback by student teachers and teachers

<table>
<thead>
<tr>
<th>Student teachers</th>
<th>Teachers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hit</td>
<td>34.3</td>
</tr>
<tr>
<td>Approach</td>
<td>33.3</td>
</tr>
<tr>
<td>Hit preparation</td>
<td>16.5</td>
</tr>
<tr>
<td>Jump</td>
<td>9.5</td>
</tr>
<tr>
<td>Global movement</td>
<td>4.8</td>
</tr>
<tr>
<td>Landing</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Analysis of the critical elements on which subjects focused provided an even more accurate picture of feedback referent. Each feedback dealing with the spike was classified in one of 14 categories of critical elements. The most important features for achievement (Kulgowczuk, 1994) were among the most frequent categories of referent (see figure 2). This finding suggests that all subjects were concerned by the main critical elements of the performance.

Three criteria considered as less important (jump height, wrist movement and lifting of the non-hitter arm) averaged at a relatively high ratio of the feedback. Easily observed, at least one of these criteria was identified among the three most frequent categories of spike referent by each subject from groups A, C and D. The experts were the only subjects who did not pay attention to these criteria. The non-specialist teachers expressed many of these criteria. They focused their feedback more frequently on 'jump height' than other subjects: 21.5 Vs 6.7% (U = 1; p = .02). Their interventions like 'You don't spring enough!', could hardly improve the learner's technical skill. Student teachers and non-specialist teachers emphasized more frequently 'wrist movement' than did the experts: 13.4 Vs 4.9% (U = 2; p = .05). Among rarely used criteria, it is surprising to observe categories like 'back movement', 'hitting arm extension', and 'feet jamming'. They are usually considered as pre-requisites to excel in spiking.
Figure 2: Ranking of the specific criteria of the spike referent feedback
Analysis of the spike referent confirmed that, when combined with teaching/coaching experience, the specialization level could improve the specificity of feedback. Experts were focusing more on important criteria than other subjects. It is noteworthy that subjects 5 and 6, the most experienced teachers in teaching/coaching volleyball showed a similar profile of spikes' referent. They centred their interventions on timing and player/ball placement. Expertise could help teachers to use the same frame of reference and select the most important criteria. Besides the quality of the performance, the specialization level and the teaching/coaching experience were two key factors in the selection of the feedback message.

Analysis of the feedback referent suggested that teachers with considerable specialization expressed frequent specific feedback. Otherwise, specialist student teachers centred their feedback on easily recognizable criteria. Teachers tended to balance their reactions between all skills involved in the teaching sequence. Furthermore, they centred their feedback on preparatory parts of the movements while student teachers were more concerned about terminal and spectacular aspects of the movements. This confirmed the results of a former study by Cloes et al. (1988). Moreover, non-specialist student teachers would focus their interventions on unimportant criteria.

4. Adequacy of feedback
Feedback was relatively appropriate (85.4%). However, compared with previous studies, the ratio of appropriate feedback was lower in this research. This difference could be explained by a high ratio of undetermined feedback (18.3%) depending on a high proportion of non-specific feedback.

Analysis of individual data showed that specialists gave more appropriate feedback than non-specialists: 70.5% vs 58.2% (U = 4; p = .013). Furthermore, specialist teachers were always among the subjects providing top quality feedback. This observation could be related to studies dealing with task analysis where specialization level was identified as a factor enabling to improve the error identification process (Girardin & Hanson, 1967; Harari & Siedentop, 1990).

It is noteworthy to observe again that subjects 5 and 6, the subjects characterized by the most expertise, were among the subjects giving the most appropriate feedback. They combined high rate and high quality feedback. Other subjects gave a high rate of feedback without reaching the same level of appropriateness. A positive influence of combining a high level of specialization and a large teaching/coaching experience was probable.

5. Discussion
Diagnosis and information processing depend on several factors, characteristics of the performance on the one hand, memory, individual variables, and proactive decisions on the other hand. In this study, experimental conditions enabled researchers to neutralize the influence of the performance. Subjects had to react to the same motor response. It could be inferred that variability of feedback was only influenced by the other factors. Specialization level and teaching/coaching experience are related to these variables. Figures 3 and 4 illustrate trends observed in results. The findings suggested the following interpretations:

- Specialization level tended to relate positively to the specificity (refer figure 3) and adequacy (refer figure 4) of feedback. This relationship would arise from an improved error identification process depending on the knowledge of the task. During practice, athletes are continuously involved in an observation process and receive feedback about the appraisals from partners, coach or other sources. Progressively, they improve their ability to detect errors.

- Teaching/coaching experience would relate to the frequency of feedback (refer figure 3). It is suggested that teachers process information quicker than student teachers; they are used to commenting upon pupils' motor performances. Novice teachers are probably able to identify some errors but are unable to react immediately; they lack the ability to link the error to the message: '... the game goes too quickly'.
Figure 3: Evaluation of specificity and frequency of feedback according to teaching experience and volleyball ability of the subjects

- When educators are not specialists, teaching/coaching experience would slightly improve the quality of criteria's selection (refer figure 4); they have not yet built their own framework and require more opportunities to develop it.
- When educators are specialists, teaching/coaching experience would substantially improve the quality of criteria's selection (refer figure 4); they choose to react on important errors due to a deeper knowledge of the skills. Such knowledge would be based on previous trial and error experiences.

As pointed out by similar patterns of experts' feedback, they could be considered as guidelines in teacher preparation. Using video-recorded motor responses as a starting point, comparison of beginners' and experts' feedback could be one way to improve the quality of that teaching/coaching skill.

Identification of classes giving similar types of feedback

Five classes of subjects were singled out by cluster analysis (refer table 4).

(i) Specialist teachers (S4, S5 and S6) were found in class 1, where the feedback characteristics seemed to be of more value to the learners' progress. One specialist student teacher (S3) was also included in that class. This subject had previous experience in volleyball coaching. Both specialization level and teaching/coaching experience were seen as playing a prominent role in the mechanisms of feedback.

(ii) Subjects from the class 2 gave non-specific feedback. Both were student teachers. A lack of specific knowledge could explain the presence of S9 in this class characterized by non-specific and undetermined feedback. S1 was the highest
Specificity

- Specialist student teachers
- Non specialist student teachers

Adequacy

- Specialist student teachers
- Non specialist student teachers

Frequency

- When educators are not specialists, teaching/coaching experience would slightly improve the quality of criteria's selection (refer figure 4); they have not yet built their own framework and require more opportunities to develop it.

- When educators are specialists, teaching/coaching experience would substantially improve the quality of criteria's selection (refer figure 4); they choose to react on important errors due to a deeper knowledge of the skills. Such knowledge would be based on previous trial and error experiences.

As pointed out by similar patterns of experts' feedback, they could be considered as guidelines in teacher preparation. Using video-recorded motor responses as a starting point, comparison of beginners' and experts' feedback could be one way to improve the quality of that teaching/coaching skill.

Identification of classes giving similar types of feedback

Five classes of subjects were singled out by cluster analysis (refer table 4).

(i) Specialist teachers (S4, S5 and S6) were found in class 1, where the feedback characteristics seemed to be of more value to the learners’ progress. One specialist student teacher (S3) was also included in that class. This subject had previous experience in volleyball coaching. Both specialization level and teaching/coaching experience were seen as playing a prominent role in the mechanisms of feedback.

(ii) Subjects from the class 2 gave non-specific feedback. Both were student teachers. A lack of specific knowledge could explain the presence of S9 in this class characterized by non-specific and undetermined feedback. S1 was the highest

Critical elements’ meaning

Table 4.
Classes of subjects identified by cluster analysis

<table>
<thead>
<tr>
<th>Class</th>
<th>Subjects</th>
<th>Feedback characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>S3 S4 S5 S6</td>
<td>Specific positive and evaluative feedback</td>
</tr>
<tr>
<td>2</td>
<td>S1 S9</td>
<td>Non-specific positive evaluative and undetermined feedback</td>
</tr>
<tr>
<td>3</td>
<td>S2 S12</td>
<td>Negative and inappropriate feedback</td>
</tr>
<tr>
<td>4</td>
<td>S10 S11</td>
<td>Non-specific negative feedback, focused on secondary aspects of the task</td>
</tr>
<tr>
<td>5</td>
<td>S7 S8</td>
<td>Few patterned feedback</td>
</tr>
</tbody>
</table>
skilled subject. Considering his background, he was expected to show a more appropriate profile of feedback. His poor performance could be related to his feelings of discomfort in the simulated teaching situations combined with too high a level of self-sufficiency.

(iii) One specialist student teacher (S2) and one non-specialist teacher (S12) formed class 3. Their feedback was quite disapproving and inappropriate. Subject 2 was the weaker specialist. Influence of the specialization level could be in conflict with internal and external factors, which impede the process of error identification. Before becoming a favoured variable to feedback, the specialization probably needs to reach a critical level. Moreover, individual characteristics could relate to feedback. Thus, both subjects of class 3 were rather critical individuals. The competency in gymnastics of the non-specialist teacher (S12) probably led him to consider the activity as a closed skill and to select criteria related to an idea model of performance.

(iv) S10 and S11, two non-specialist teachers, focused their feedback on secondary aspects of the task (class 4). Without any specific practical framework, most of the non-specialist teachers were unable to manage their own teaching/coaching experience in order to select appropriate information.

(v) Most non-specialist student teachers (S7 and S8) seemed to lack a specific pattern of feedback (class 5). Without any framework, they did not follow a feedback strategy.

Conclusions
The purpose of the study was to identify the inter-individual variability of feedback emitted by physical educators characterized by different specialization levels and teaching/coaching experience. The originality of this study dealt with the analysis of the feedback variability in a simulated teaching setting. This research technique enabled the subjects to react to identical motor performances and to be involved in standardized conditions.

Analysis of the characteristics of feedback emitted by different groups of subjects and the cluster analysis led to parallel results. At different degrees, the specialization level and the teaching/coaching experience were related to several feedback characteristics. Theoretical and practical knowledge arising from specialization would improve the accuracy of the diagnosis by identifying relevant cues. Teaching/coaching experience would enhance the delivery of the message in easing the information processing.

Studying feedback in simulated teaching conditions of experts and novice physical educators could be a base from which to develop training procedures leading to improvement of the whole feedback process.

References


Résumé

Variabilité interindividuelle des feedback des enseignants. Etude en situation d’enseignement simulé

Dans l’enseignement des activités physiques, les réactions de l’enseignant à la prestation des élèves (feedback) constituent l’une des variables déterminantes dans la réussite pédagogique. Les réactions à la prestation se caractérisent par une grande variabilité interindividuelle. Leur émission fait suite à un processus comportant deux phases: le diagnostic et la prescription.

Afin de mettre en évidence le rôle exercé par les facteurs liés aux caractéristiques de l’enseignant sur celles de leurs interventions, nous avons comparé les réactions à la prestation émises par 12 éducateurs physiques lors d’une séance d’enseignement simulé. Grâce à cette technique vidéo, les sujets réagissent à partir de performances motrices identiques. Les feedback ont été comparés entre quatre groupes de sujets caractérisés par des niveaux d’expérience sportive et pédagogique similaires. Une analyse factorielle par correspondance a permis d’identifier des classes de sujets émettant des feedback de même type.

Les résultats tendent à souligner que (1) le niveau d’expertise dans la pratique de l’activité enseignée exerce une influence sur la première de ces phases; (2) le degré d’expertise pédagogique contribue à améliorer l’émission du message.
Zusammenfassung
Individuelle Unterschiede des Lehrer-Feedbacks — eine Studie über simulierte Lehrbedingungen

Gudrun Doll-Tepper

Resumen
Variabilidad interindividual para la administración de retroalimentación por parte de los profesores. Estudio basado en condiciones de enseñanza simuladas
En la enseñanza de la educación física, la retroalimentación que el profesor administra a los alumnos está considerada como una variable de eficiencia. Por otro lado, la retroalimentación está caracterizada por una amplia variabilidad entre individuos y su emisión requiere determinadas técnicas para su diagnóstico y su comunicación.

El propósito de este estudio se centró en comparar la retroalimentación administrada por 12 educadores físicos a lo largo de una sesión de enseñanza simulada con el objetivo de dar forma a las relaciones existentes entre características de los profesores y la retroalimentación. A través de una técnica con video, los sujetos administraron feedback sobre rendimientos motores idénticos. Dicha retroalimentación fue comparada entre cuatro grupo de sujetos caracterizados por tener niveles similares de especialización y experiencia pedagógica (métodos estadísticos no paramétricos). El análisis de cluster permitió la identificación de las clases de sujetos en función del tipo de retroalimentación administrada.

Los resultados demostraron: En primer lugar, que la especialización se identificó como una variable que influenciaba la calidad del diagnóstico. Y en segundo lugar, que la experiencia de enseñanza era una variable que realzaba la calidad del mensaje dirigido.

José Campos Granell