The purpose of this study was to determine whether varying levels of expertise would produce differences in selected indicators of effective teaching performance. Eighteen elementary physical education teachers were grouped based on stages of pedagogical expertise development as suggested by Berliner (1988). The three groups included novice/advanced beginner teachers (1-2 years of experience), competent teachers (3-4 years), and proficient/expert teachers (5-8 years). In addition, three students were randomly selected in each teacher's class. Data were collected on the percent of transition time, subject matter motor (SMM) time, Academic Learning Time-Physical Education (ALT-PE), students' off-task behavior, and negative feedback. Analysis of variance procedures resulted in no statistically significant differences between teacher groups on any of the selected indicators, pointing to similar levels of teaching performance across groups. The majority of this research concentrated on variables in the cognitive domain; therefore, it is suggested that further efforts be employed using more subtle behavioral indicators. (Author/AMH)
Novice and Expert Physical Education Teachers:
They May Think and Decide Differently . . .
But Do They Behave Differently??
Hans van der Mars, E. William Vogler, Paul W. Darst
Arizona State University
Barbara Cusimano
Oregon State University
Novice and Expert Physical Education Teachers:
They Think and Decide Differently . . .
But Do They Behave Differently??

Abstract
The purpose of this study is to determine whether varying levels of expertise would produce differences in selected indicators of effective teaching performance. Eighteen teachers were grouped based on stages of pedagogical expertise development suggested by Berliner (1988). The three groups included Novice/Advanced Beginner teachers (1-2 years of experience), Competent teachers (3-4 years), and Proficient/Expert teachers (5-8 years). In addition, three students were randomly selected in each teacher’s class. Data were collected on the percent of transition time, subject matter motor (SMM) time, ALT-PE, students’ off-task behavior, and negative feedback. Analysis of Variance procedures resulted in no statistically significant differences between teacher groups on any of the selected indicators, indicating similar levels of teaching performance across groups.

Key words: Teaching expertise, teaching behavior, Academic Learning Time-Physical Education
The study of expertise in teaching is receiving increasing attention, and its knowledge base is expanding rapidly. Most research in this area is rooted in cognitive developmental psychology, and has used a design in which comparisons are made between novices (e.g. student teachers) and experts (e.g. veteran teachers with more years of teaching experience). Initial classroom studies indicate that compared to novices, expert teachers interpret and solve problems differently, recognize classroom events with greater speed and accuracy, and actively plan and teach their students certain organizational and managerial routines in the beginning of the school year (Berliner, 1986, 1988; Brooks & Hawke, 1987; Carter, Cushing, Sabers, Stein, & Berliner, 1988; Leinhardt & Greeno, 1986).

In physical education the knowledge base of teachers' cognitive processes within the expertise realm is growing rapidly. Research on preactive (i.e., planning) decision patterns of expert and novice physical educators has shown that expert teachers make more preactive decisions and were more likely to use equipment for other than its traditional purposes than novices. In addition, they have a broader knowledge base, utilize implicit theories of instruction, consider more alternatives (i.e. adaptations) for possible changes that might be needed, make more use of retrieval systems from their memory base, and do not need as much time to plan as do novices. Novices also reported higher levels of anxiety and stress when
planning (Housner & Griffey, 1985; Howell, 1987; Sherman, 1979, 1983; Sherman, Sipp, & Taheri, 1987).

Studies on differences in interactive (i.e. in-class) attention and decision making patterns have shown that experienced physical educators focus more on students' performance and involvement (Housner & Griffey, 1985). In addition, they made decisions that were mostly managerial in nature, followed by activity decisions. Novices on the other hand made more activity-based decisions. Experts were more likely to make "in-flight" adjustments in their original plans where necessary and did not "push the panic button" when problems occurred (Sherman, 1983). As with classroom teachers, experts put greater emphasis on activities in the beginning of the school year aimed at developing managerial routines and procedures (Nelson, 1988).

Expert physical educators were found to have similar thought processes and decision making patterns as those found in expert classroom teachers. Specifically, they more accurately interpreted instructional events, and solved problems that arose during class more creatively (Ashy, Howell, & Lee, 1989; Howell & Lee, 1988; Nelson, 1988). Ashy et al. (1989) noted that where experts were willing to make changes in their plans while teaching, novices appeared unwilling to make adjustments for fear of losing control.

Expert-novice research in physical education has also
focused on teachers' knowledge structures. This research is based primarily on findings in other fields (e.g. chess, and medical diagnosis) where experts have been shown to have not only more knowledge, but they also appear to have it organized in more meaningful "chunks". Hacker (1989) replicated these results on knowledge structures specific to classroom management between experienced and novice physical educators. Lynn, French, Rink, Lee, and Solmon (1990) used a modified "ordered tree" method to study differences in overall pedagogical knowledge structures of expert and novice physical educators. They found that experts have "Knowledge that they chunk into meaningful units that make sense out of their experiences, and, that they establish logical relationships between those chunks of information" (p. 10).

Research on pedagogical expertise has concentrated primarily on the study of teachers' cognitive processes. Recently, Siedentop and Eldar (1989) provided a behavioral perspective on the nature of pedagogical expertise. It was argued that expertise is highly context- and subject matter bound; lies "at the nexus of skillful teaching and thorough command of the subject matter." (p. 257); and is primarily performance based. Pedagogical expertise from a behavioral perspective was explained in terms of developing stimulus control where teachers have developed fine-tuned discrimination skills, specific to particular areas of the subject matter, allowing them to recognize relevant class events more quickly and thus react to
them faster. If expertise lies partly in highly skillful teaching performance that develops over long periods of time, one would expect to see behavioral differences emerge time between more and less experienced teachers, assuming, of course, that skillful teaching, over time, is differentially reinforced.

Most research on pedagogical expertise has concentrated on comparing thinking and decision making processes of teachers who are in the first (novice) and last stage (expert). However, little is known about if/how expertise in teaching takes shape (in physical education), either from a longitudinal or cross-sectional perspective (Borko & Livingston, 1989). Nor is there much evidence on expertise from a behavioral perspective. Using Berliner's (1988) proposed model of pedagogical expertise development, this study sought to determine whether differing levels of expertise would be reflected in selected indicators of effective teaching.

Methods

Subjects and Settings

Teachers. Eighteen certified elementary physical educators from three different school districts participated in the study following standard informed consent procedures. Based on their years of teaching experience and related indicators, teachers were grouped in one of three expertise level categories as proposed by Berliner (1988) (see Table 1). Teachers in the Novice/Advanced beginner group were selected purely on the number
of years of teaching experience. In addition to the greater number of years of experience indicating passage through the induction stage, teachers in the Competent group had been teaching at the same school for at least three years.

Proficient/Expert group teachers included professionals who had demonstrated expertise in a number of ways: (a) Selected as the state's physical education teacher of the year; (b) served as school district physical education supervisors/mentors; (c) served as state officers in physical education professional organization, including presidency; (d) presented workshops at both national, regional and local level conferences; (e) involved in teaching research projects with a local university; and (f) described as exemplary practitioners by principals.

Table 1 about here

Teachers utilized the same curriculum approach (Dauer & Pangrazi, 1986). While the selection of activities could be not controlled, the class format was the same across all teachers, including an (a) introductory activity, (b) fitness development, (c) lesson focus which was aimed at the practice of skills, and (d) game portion. Analysis of Variance (ANOVA) procedures indicated no significant difference between groups indicating teachers' similar time allotment to each lesson section.

Students. Three students in each of the aforementioned teachers' classes were randomly selected to serve as subjects.
There were 27 boys and 27 girls, including 18 first graders, 6 second graders, 6 third graders, 12 fourth graders, 9 fifth graders, and 3 sixth graders. They were distributed evenly across the three groups. The student sample included youngsters of Anglo-Saxon, Native American, Asian, Hispanic, and Afro-American heritage.

Setting. The schools where data were collected were all located in middle class suburban areas. Class sizes observed ranged from 24 to 31 students. At each school manipulative equipment (i.e., balls, hockey sticks, hoops, jumpropes etc.) was available for each individual student, and all but two classes were taught indoors.

Procedures

Teachers and target students were videotaped once during a regular 30 minute class. A character generator stopwatch projected the elapsed class time on each videotape. Each teacher wore a wireless microphone to capture all verbal behavior which was dubbed onto the videotape record. Two video cameras were used to ensure that the obtained records of subjects' behavior would be complete.

Dependent Variables

Videotapes were analyzed to collect data on five dependent variables each of which has been shown to be indicative of effective teaching practices (i.e., affect student achievement). First, percent of total feedback that is negative, was used as an
indicator of the overall learning climate. Soar and Soar (1979) noted a number of large scale correlational studies indicating that absence of negative affect is critical if student learning is to occur.

Second, percent of class time allocated to Subject Matter Motor (SMM) activities and percent of transition time were included. Both are indicators of the teachers' organizational and managerial ability to put students in contact with relevant content. Brophy and Good (1986) and Berliner (1979) have reported a positive link between student achievement and time allocated to subject matter content. Time devoted to non-instructional or organizational activities (i.e., Transition time) detracts from or limits the functional time spent in contact with such content (Berliner, 1979).

Third, percent of students' successful task engagement, often referred to as Academic Learning Time-Physical Education (ALT-PE) or Opportunity to Respond was selected. Given the variability in the nature of the activities taught we opted to use ALT-PE as the indicator of successful motor engagement. There is considerable support for using this variable as an indicator of student learning both in classroom (Fisher, Berliner, Filby, Marliave, Cahen, & Dishaw, 1980) and physical education (Metzler, 1988) settings.

Finally, percent of students' off-task behavior was selected as an indicator of student behavior management. While there no
direct link between off-task behavior and student achievement (Siedentop, 1983), it is generally accepted that for students to learn a minimum level of orderliness is required for subject matter related instruction to occur (Doyle, 1986).

Data Collection.

Teachers' verbal negative feedback data were collected using standard event recording. The remaining data were collected using the interval recording version of the Academic Learning Time-Physical Education instrument (Parker, 1989) with noncontinuous observation control (6 secs. observe/6 secs. record) (Tawney & Gast, 1984).

Data Accuracy

Traditionally, observer reliability is represented by the percentage of inter-observer-agreement (IOA). However, as Cooper, Heron and Heward (1987) indicated, "High interobserver agreement (e.g. 90%), does not necessarily mean that observations were accurate." (p. 93). In the present study, the "true" dimensions of all process behaviors of interest were permanently captured on videotape. As a result, accuracy checks on all data samples could be completed, and IOA calculations were not necessary (Johnston & Pennypacker, 1980).

Four videotapes (22%) were randomly selected to determine data accuracy. Data accuracy was checked by having a trained observer compare the observed values on the coding sheets with the actual behavior patterns on the videotape record across
intervals. Instances where values recorded on the coding sheets matched the actual (i.e. true) behaviors on the videotape were counted as an instance of accuracy. Those instances where the two did not match were counted as an inaccuracy. A data accuracy percentage was then calculated by using an equation identical to the one used for determining IOA (i.e., \([\text{Accuracy}/\text{Accuracy} + \text{Inaccuracy}] \times 100\)). Data accuracy percentages are presented in Table 2. Prior to data analysis, any discrepancies noted were reviewed by one of the investigators with the accuracy reviewer and through error resolution collected data were edited, i.e. changed to agreed upon behavior percentage(s).

Table 2 about here

**Data Analysis**

Following data collection and accuracy checks, raw data for ALT-PE and off-task behavior data were averaged within class. These were then averaged again within each teacher group of expertise. Statistical analysis was performed using Analysis of Variance (ANOVA) on each of the five dependent variables.

**Results**

Means and standard deviations per teacher group across dependent variables are presented in Table 3. The following data patterns across groups emerged: Time available to be actively engaged in motor tasks (SMM Time) was highest for students of
teachers in the Competent group, while similar for the Novice/advanced and Proficient/Expert group. Transition time was similar for the Novice/advanced beginner and Competent groups, and diminished in the Proficient/Expert group. The percent of feedback that was negative in nature was low across across groups, showing a small increase with teachers in the Proficient/expert group.

Table 3 about here

Successful motor engagement (i.e., ALT-PE) levels across groups were higher on the average than those reported in previous descriptive studies, and increased gradually across groups. The ratio of SMM Time and ALT-PE is an indicator of (a) how well activities were organized in terms of providing students with the opportunity to be actively engaged (instead of being forced to wait in lines for example) and (b) the appropriateness of task selection by teachers in terms of difficulty level. A higher ratio would be indicative of less effective activity organization and/or less appropriate task selection. In this study the ratio decreased gradually with increasing levels of expertise 1.7 : 1 to 1.6 : 1 to 1.3 : 1. Finally, students' off-task levels were very low in each group, indicating little change across groups. ANOVAs indicated no statistically significant differences between groups (see Table 3).
Discussion

The data reported here across each of the selected indicators of effective teaching (i.e. classroom management, behavior management, emotional climate, successful motor engagement) indicated that teachers with more advanced levels of expertise are not necessarily more effective than those at initial levels of expertise. That is, their thought processes and patterns of decision making might have been different, but through the "windows" of the selected effective teaching indicators used here they did not look any different.

There are two plausible explanations for the lack of differences between groups. The sample of teachers may not have been a true cross-sectional representation of teacher expertise and the selected dependent variables may simply have been a measure of effectiveness rather than expertise.

First, our novice group may have been atypically effective. For example, ALT-PE levels were either similar or higher than those reported in previous studies (e.g., Dodds, Rife & Metzler, 1982). It is possible that those teachers in the novice/advanced beginner group were not really behaving as novices. For example, in review of their videotapes, it was noticed that they made frequent use of selected organizational routines and procedures typically reflective of teachers with more expertise (Brooks & Hawke, 1987; Nelson, 1988). That is, they had developed a repertoire of instructional skills that enabled them to reach the
performance level of their more experienced colleagues. Thus they had reached what Siedentop and Eldar (1989) referred to as "experienced effectiveness" earlier than one might reasonably expect.

Second, the dependent variables selected for this study may not have been sensitive enough to detect differences between levels of expertise. Assuming that expertise lies beyond effectiveness, other more fine-grained, molecular behavioral variables can/should be included in subsequent analyses. For example, ample study has been made of how time is being spent by teachers and students in physical education. Such data have typically been reported in total percentages. However, in sport pedagogy little attention has been given to other dimensional quantities of behavior or events such as transitions, including frequency/rate of transitions, or the interresponse time of transitions. For example, Arlin (1979) showed how teacher initiated transitions can disrupt the timeflow during class. Within the pedagogical expertise research paradigm a viable question would be to determine if expert teachers differ from novices (i.e., become more skilled) in both the structuring, timing and pacing of transitions and motor task activities.

Other related variables that may show developmental patterns of expertise over time could be those developed by Kounin (1970), including momentum, smoothness, and accountability. Nelson (1988) reported that expert teachers use student skill
performance to make changes in the lesson. With such attention to performance, do experts set up accountability systems in different ways than do more novice teachers?

Hawkins and Landin (1989) have started the use of field systems analysis to study experts' teaching behavior patterns. In this approach to the study of teaching environments, particular attention is given to "the contextual and temporal relationships among setting variables" (Sharpe & Hawkins, 1990, p. 4). As such, the analysis of teachers' and students' in-class behavior patterns from an (inter)behavioral perspective has by no means been exhausted and undoubtedly will provide further insight into expert teachers' handling of the daily teaching task and how it differs from approaches used by more novice teachers.

In summary, different levels of expertise, based on a tentative model of pedagogical expertise development, were not reflected in the in-class behavior of either teachers or their students. While the majority of this research has concentrated on variables in the cognitive domain, it is suggested that further efforts be employed using more subtle behavioral indicators.
References


Abstracts of research papers (p. 106). Reston, VA: AAHPERD.


Lynn, S., French, K., Rink, J., Lee, A., & Solmon, M. (1990, April). Comparison of pedagogical knowledge structures of


Table 1. **Teachers and Students Background Information.**

<table>
<thead>
<tr>
<th>Teacher group</th>
<th>n</th>
<th>Age (SD)</th>
<th>Experience (SD)</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice/Advanced Beginner (1-2 yrs)</td>
<td>7</td>
<td>26.7(4.2)</td>
<td>1.7(.5)</td>
<td>10</td>
</tr>
<tr>
<td>Competent (3-4 yrs)</td>
<td>5</td>
<td>30.2(3.7)</td>
<td>3.4(.5)</td>
<td>8</td>
</tr>
<tr>
<td>Proficient/Expert (5-8 yrs)</td>
<td>6</td>
<td>29.6(1.6)</td>
<td>7.0(1.1)</td>
<td>9</td>
</tr>
</tbody>
</table>


Table 2. **Accuracy Percentage Means and Ranges.**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Accuracy Percentage Mean</th>
<th>Accuracy Percentage Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject matter motor time</td>
<td>100</td>
<td>n.a.</td>
</tr>
<tr>
<td>Transition time</td>
<td>100</td>
<td>n.a.</td>
</tr>
<tr>
<td>Negative feedback</td>
<td>94</td>
<td>90-96</td>
</tr>
<tr>
<td>ALT-PE</td>
<td>99</td>
<td>97-100</td>
</tr>
<tr>
<td>Off-task</td>
<td>96</td>
<td>94-99</td>
</tr>
</tbody>
</table>
Table 3. **Means, Standard Deviations and ANOVA Results Across Levels of Expertise Groups.**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Novice/Adv. Beginner (SD)</th>
<th>Competent (SD)</th>
<th>Proficient/Expert (SD)</th>
<th>F</th>
<th>Ratio</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMM Time (%)</td>
<td>44.2(7.0)</td>
<td>53.0(13.4)</td>
<td>46.3(18.5)</td>
<td>.63</td>
<td>.54</td>
<td></td>
</tr>
<tr>
<td>Transition (%)</td>
<td>28.2(6.7)</td>
<td>28.3(7.8)</td>
<td>21.4(3.0)</td>
<td>2.49</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>Neg. Feedback (%)</td>
<td>7.6(3.7)</td>
<td>7.8(8.9)</td>
<td>10.5(6.3)</td>
<td>.38</td>
<td>.68</td>
<td></td>
</tr>
<tr>
<td>ALT-PE (%)</td>
<td>26.0(9.8)</td>
<td>31.9(6.5)</td>
<td>35.3(9.0)</td>
<td>1.89</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td>Off-task (%)</td>
<td>3.4(3.1)</td>
<td>6.7(2.1)</td>
<td>4.4(1.6)</td>
<td>2.74</td>
<td>.09</td>
<td></td>
</tr>
</tbody>
</table>

df: 2,17
Critical F-value: 3.59 (p > .05)