

2001

Authentic assessment in physical education: Prevalence of use and perceived impact on students' self-concept, motivation, and skill achievement

Joseph Kwame Mintah
University of Northern Iowa

Let us know how access to this document benefits you

Copyright ©2001 Joseph Kwame Mintah

Follow this and additional works at: <https://scholarworks.uni.edu/etd>



Part of the [Health and Physical Education Commons](#)

Recommended Citation

Mintah, Joseph Kwame, "Authentic assessment in physical education: Prevalence of use and perceived impact on students' self-concept, motivation, and skill achievement" (2001). *Dissertations and Theses @ UNI*. 723.

<https://scholarworks.uni.edu/etd/723>

This Open Access Dissertation is brought to you for free and open access by the Student Work at UNI ScholarWorks. It has been accepted for inclusion in Dissertations and Theses @ UNI by an authorized administrator of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.

Offensive Materials Statement: Materials located in UNI ScholarWorks come from a broad range of sources and time periods. Some of these materials may contain offensive stereotypes, ideas, visuals, or language.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

AUTHENTIC ASSESSMENT IN PHYSICAL EDUCATION: PREVALENCE OF
USE AND PERCEIVED IMPACT ON STUDENTS' SELF-CONCEPT,
MOTIVATION, AND SKILL ACHIEVEMENT

A Dissertation

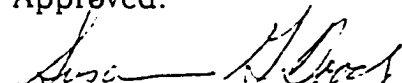
Submitted

In Partial Fulfilment

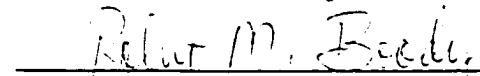
of the Requirements for the Degree

Doctor of Education

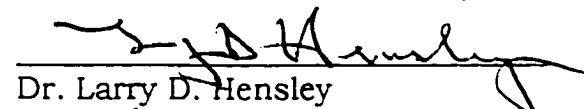
Approved:



Dr. Susann G. Doody, Co-Chair



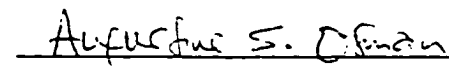
Dr. Robert M. Boody, Co-Chair



Dr. Larry D. Hensley



Dr. Linda M. Fitzgerald



Dr. Augustine S. Osman

Joseph Kwame Mintah

University of Northern Iowa

May 2001

UMI Number: 3005666

UMI[®]

UMI Microform 3005666

Copyright 2001 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Copyright by
JOSEPH KWAME MINTAH
May 2001
All Rights Reserved

AUTHENTIC ASSESSMENT IN PHYSICAL EDUCATION: PREVALENCE OF
USE AND PERCEIVED IMPACT ON STUDENTS' SELF-CONCEPT,
MOTIVATION, AND SKILL ACHIEVEMENT

An Abstract of a Dissertation

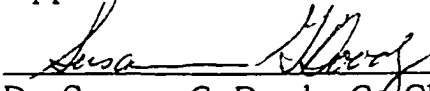
Submitted

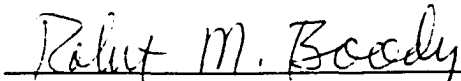
In Partial Fulfilment

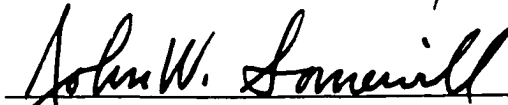
of the Requirements for the Degree

Doctor of Education

Approved:


Dr. Susann G. Doody, Co-Chair


Dr. Robert M. Boody, Co-Chair


Dr. John W. Somervill
Dean of the Graduate College

Joseph Kwame Mintah

University of Northern Iowa

May 2001

ABSTRACT

The volume of traditional assessments in the public schools of the United States has increased. At the same time, complaints about the use of standardized tests have become louder. The complaints and dissatisfaction have led to a paradigm shift in assessment. Available literature supports the use of alternative assessment. The purpose of this study was to describe and analyze the type and the extent of authentic assessment used in public school physical education. A second purpose was to investigate physical education teachers' perceptions about the impact of authentic assessment use on students' self-concept, motivation, and skill achievement.

The teachers rated use of 15 authentic assessment techniques on a scale (1 = never used through 5 = always used). In addition, they indicated their perceptions about the impact authentic assessment use has made on their students' self-concept, motivation, and skill achievement on a Likert scale (2 = strongly disagree through 5 = strongly agree). Furthermore, 12 of the teachers who participated in the mail survey were interviewed.

Two hundred and ten Iowa public school physical education teachers participated in this study. Results showed that 158 (75.2%) teachers use authentic assessment techniques. Frequency statistics indicated that teacher observation, self-observation, checklists,

observation, and event task were among the top six authentic assessments used. In addition, the perception of the physical education teachers was that authentic assessment use has had a positive impact on students' self-concept, motivation, and skill achievement.

In summary, Iowa public school physical education teachers use authentic assessment extensively. The physical education teachers perceived authentic assessment use has enhanced students' self-concept, motivation, and skill achievement positively.

DEDICATION

To my grandmother and parents, whose deep commitment to western education laid the foundation for this advanced degree. To my wife, Mrs. Olive Raleigh Mintah, and the children who encouraged me during difficult times. To my professors, friends, and colleagues who shared their time, energy, and knowledge with me, may God bless you all.

ACKNOWLEDGEMENTS

The great Dean of St. Paul's Cathedral in the seventeenth century wrote: "No man is an Iland, intire of it selfe; every man is a piece of the Continent, a part of the maine...." Today, I write to express my profound gratitude to those who have contributed in the making of this dissertation, known and unknown, friends, colleagues, and gentle critics.

Special thanks go to Dr. Susann G. Doody, my program adviser, who after retiring from active teaching agreed to stay and chair this project. I am also indebted to Dr. Robert M. Boody, co-chair, Dr. Larry D. Hensley, Dr. Augustine S. Osman, and Dr. Linda M. Fitzgerald, all committee members. Without your directions, this dissertation would not have been possible.

I wish to express my deep gratefulness to Dr. Christopher R. Edginton for willingly allowing me to use the facilities in the school of Health, Physical Education and Leisure Services. I also thank Dr. Sharon Huddleston for allowing me to use the psychomotor lab during the research process.

To my wife Olive and children, who for months have endured with me the stress of writing a dissertation, I am forever appreciative. Finally, I recognize my dependence on my Lord and Savior Jesus Christ, who is consistently consistent even in the face of my inconsistency.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION.....	1
Rationale of the Study.....	4
Statement of the Problem.....	7
Research Questions.....	8
Limitations.....	8
Delimitations.....	9
Assumptions.....	9
Definition of Terms.....	9
II. REVIEW OF LITERATURE	11
Authentic Assessment in General Use.....	12
Authentic Assessment Defined.....	13
Characteristics of Authentic Assessment.....	14
Authentic Assessment Uses in the Public School.....	17
The Case for Authentic Assessment	20
Authentic Assessment Uses in Physical Education	24
Portfolio	28
Exhibitions/Demonstrations	30
Oral Discourse	31
Written Essay	31

CHAPTER	PAGE
Constructivism	32
Applying Constructivism in Physical Education	36
Summary	37
III. METHODS	39
Design	39
Subjects	40
Sampling	40
Demographic Information	40
Dependent Measure	41
Physical Education Authentic Assessment Inventory (PEAAI)	41
Pilot Testing	42
Section I: Participating Pilot Test	42
Section II: Undeclared Pilot Test	43
Procedure	44
Data Gathering	44
Phase I: Mail Survey	44
Phase II: Interview	45
Description of Data	46
Scoring	46
Data Analysis	47

CHAPTER	PAGE
Mail Survey	47
IV. RESULTS	49
Subjects Demographics	49
Extent of Authentic Assessment Use	50
Reported Use of Individual Authentic Assessment Techniques Among "Users" by Gender	51
Reported Use of Individual Authentic Assessment Techniques Among "Users" by Grade Level	54
Reported Frequency of Use of Popular Authentic Assessment Techniques by Grade Level	56
Teacher Observation	57
Self Observation	58
Demonstration	58
Peer Observation	58
Event Task	59
Checklists	59
Reported Frequency of Use of Popular Authentic Assessment Techniques by Gender	60
Perceived Impact of Authentic Assessment Use on Students' Self-Concept, Motivation, and Skill Achievement	61
Self-Concept	62
Motivation	62
Skill Achievement	62

CHAPTER	PAGE
Summary	63
Differences in Perceptions About the Impact of Authentic Assessment Use on Students' Self-Concept, Motivation, and Skill Achievement	63
Interviews	65
Participants	65
Protocol	65
Questions Guiding the Interview	66
Data Collection	67
Data Analysis	68
Emergent Themes: Teachers' Perceptions	68
The Theme of Motivation, Self-concept, and Skill Achievement	69
The Theme of Effort and Participation as Assessment Techniques in Physical Education	72
The Theme of Fitness and Skill Testing in Physical Education	73
The Theme of Paper and Pencil Testing	75
Summary	75
V. DISCUSSION AND RECOMMENDATIONS	77
Discussion	77
Extent of Authentic Assessment Use	78
Use of Authentic Assessment Techniques among "Users"	80

CHAPTER	PAGE
Non-"Users" of Authentic Assessment Techniques ...	83
Perceived Impact About Authentic Assessment Use on Students' Self-concept, Motivation, and Skill Achievement	85
Recommendations for Research	86
Recommendations for Practice	87
REFERENCES	88
APPENDIX A	96
APPENDIX B	97
APPENDIX C	99
Section A: Authentic Assessment Techniques	99
Section B: Physical Education Teachers' Perceptions	101
APPENDIX D	104

LIST OF TABLES

TABLE	PAGE
1. Reported Use of Individual Authentic Assessment Techniques Among "Users" by Gender	53
2. Reported Use of Individual Authentic Assessment Techniques Among "Users" by Grade Level	55
3. Means and Standard Deviations of Frequency of Use of the Popular Authentic Assessment Techniques by Grade Level	57
4. Means and Standard Deviations of Frequency of Use of the Popular Authentic Assessment Techniques by Gender	60
5. Means and Standard Deviations of Perceived Impact of Authentic Assessment Use on Self-Concept, Motivation, and Skill Achievement by Gender by Grade Level	64

CHAPTER I

INTRODUCTION

When an educational problem persists despite the well-intentioned efforts of many people to solve it, it's a safe bet that the problem hasn't been properly framed. Assessment in education has clearly become such a problem . . . (Wiggins, 1989b, p. 703).

In 1995, when the educational reform movement led to new educational standards, school districts were asked to find ways to improve students' achievement (Cleland & Stevenson, 1997). Among the noticeable changes that accompanied the educational reform movement was the move toward performance-based assessment. As a result, school districts have experimented with new assessment formats (Zhu, 1997). Today, in the public schools, authentic assessment has caught the attention of many educators (Powell, 1993).

Authentic assessment uses multiple scoring systems to measure students' habits and repertoires on significant tasks related to life outside the classroom (Wiggins, 1989b; Worthen, 1993). According to Wiggins (1989b), "authentic assessment replicates the challenges and standards of performance that typically face writers, business people, scientists, community leaders, designers, or historians" (p. 705). In addition, authentic assessment enables educators to watch a learner pose and tackle problems, arrange arguments, marshal evidences, and take purposeful actions to address and solve ambiguous problems. With

authentic assessment, students' competence is not assessed from one performance but through a series of activities (Lund, 1997). Students are exposed to different assessment tasks so that they can demonstrate their competence. The assessment tasks have contextual significance (Hensley, 1997; Wiggins, 1989b), and are directed at the behavior, knowledge, or feelings the teacher wishes to measure. Authentic assessment, therefore, focuses on the product and the quality of performance. With authentic assessment, students are more actively involved in the learning process. In addition, students know how they will be evaluated ahead of the actual assessment, which often results in higher levels of students' interest and motivation (Mitchell, 1992).

Many types of assessment are reported in the literature for physical education, including written essays, oral discourses, exhibitions and event tasks, portfolios, checklists, report cards, student logs, peer observation, self observation, and parental report (Lund, 1997; Parker, 1998). Written essays are used to describe and explain facts in context. Written essays enable students to use analysis, synthesis, and critical thinking skills to present materials logically. Oral discourses or interviews give students the chance to show their competence and knowledge. For example, a student might discuss the merits of a zone versus player marking defense during an oral discourse. Through oral discourse, students synthesize knowledge, draw conclusions, make

decisions, and justify their choices. Exhibitions and event tasks are other authentic assessment practices used in physical education. Exhibitions are extensive demonstrations of a student's skills or performance (Feuer & Fulton, 1993). Event tasks, on the other hand, are exhibitions that can be done in one class period (Lund, 1997). Another type of authentic assessment measure use in physical education is the portfolio. Portfolios are collections of students' work over a period (Melograno, 1994^{or 1999}2000), and are superb tools for documenting students' gains (Jones, 1993; Kirk, 1997; Ryan & Miyasaka, 1995). Portfolios provide students with the opportunity to explore goals (Kirk, 1997), and they can be employed to whatever purpose necessary because they are very flexible (Hauge, 1997).

Authentic assessment is an assessment approach that follows from the constructivist paradigm of teaching and learning. Constructivist learning is based on the idea that children's minds are not blank slates (Von Glasersfeld, 1990). Students have a set of beliefs, theories and perceptions. Learning happens when these are challenged through conversation, hands-on activity or experience (Noel, 1993). In constructivism, the learner as a whole person is the focus. Appropriate assessment, according to constructivist learning theorists, will consider individual differences in experience and ability that focus on providing

assessment on an individual ongoing basis (DeVries, 1987; Von Glasersfeld, 1990).

Rationale of the Study

The move toward the use of authentic assessment emanated from the view that most traditional assessments are not good representations of subject matter problems and challenges or the students' actual/usable knowledge (Dana & Tippins, 1993). Current tests do not tap many skills and abilities that students need to develop to be successful in later life and schooling. According to Wiggins (1993), most traditional assessments measure common and narrowly defined knowledge that is incompatible with the aim of any robust education for lifelong learning. Furthermore, bona fide intellectual performance is inherently personalized. And, because the meanings, strengths and aspirations derived from education are intrinsically idiosyncratic, using traditional assessment short-circuits the vital educational exchange between individual and meaning (Wiggins, 1993).

Another force behind the development of authentic assessment originates from new views about teaching and learning. Katims, Nash, and Tocci (1993) found recent increases in the emphasis of connection within and across disciplines. To them, such blurs of boundaries between subject-matter categories have made learners more active, collaborative and higher-order thinkers. With authentic assessment,

instruction and assessment are interlaced (Diez & Moon, 1992; Head, 1996; Lund, 1997; Wiggins, 1989a). Assessment requires students to apply many skills acquired in class and to use these skills as foundations to build upon them. For assessment to measure high-level knowledge, decide the extent to which students have achieved the goals of a class, and serve as one important tool teachers can use to motivate students and facilitate learning, Madaus and Kellaghan (1993) recommended a shift toward authentic assessment.

Today's education demands that students do more than memorize information used to solve problems (O'Neil, 1992). With the rapidly changing educational landscape, traditional psychometric assessment tools, such as multiple-choice tests and sport skill tests, may no longer be adequate for assessing learning outcomes (Wood, 1996). To adequately assess students' mastery, assessment should require a meaningful task designed to be representative in the field (Lund, 1997).

In addition, a big variance in growth and experience exists among young adolescents in the schools. For example, in the middle/junior high school levels, students differ in size and strength. Most traditional assessments in physical education, such as fitness testing, neglect this great variability and diversity among these students (Kritt, 1993). Because traditional physical education assessments do not consider individual differences among young adolescents, some students find it

difficult to keep pace with the classroom activities. For those students, their motivation to learn diminishes despite the need to compete for better grades. To educate students with varying physical, cognitive and social-emotional development as well as increase students' self-concept, motivation, and skill achievement physical educators should use different assessment formats.

Although many agree that authentic assessment increases students' self-concept, motivation, and skill achievement some differences in opinions exist in the literature about its use (Cizek, 1991; Meyer, 1992). For example, the lack of psychometric data about authentic assessment tasks has created disagreement among physical educators about the use of authentic assessment (Herman & Winters, 1994; Lund, 1997; Worthen, 1993). In addition, the subjective nature and the lack of acceptable criteria for comparing authentic assessment measures in physical education have created ambiguity when used to make critical decisions.

In physical education, there is confusion about the meaning of the term authentic assessment. Herman, Aschbacher, and Winters (1992), acknowledged that performance assessment, alternative assessment, and authentic assessment are synonymous, whereas Marzano, Pickering, and McTinghe (1993) differentiated among the three. Such discrepancies confuse physical educators about the ideas of performance assessment

equating that with assessment of performance. Furthermore, authentic assessment requires time to plan and evaluate (Lund, 1997). The time consuming nature of authentic assessment tasks has made some physical educators think it is labor intensive. Finally, physical education requires extensive demonstration of students' knowledge. Using traditional psychometric assessment techniques in physical performance lowers students' motivation and self-concept. To make physical education class real, students must be provided with opportunities to show their expertise in a variety of situations such as portfolios, exhibitions or demonstrations (Weinberg, 1996).

Statement of the Problem

In physical education, there is disagreement, or at the least confusion, about the use of the term authentic assessment (Herman et al., 1997; Marzano et al., 1993). In addition, there is a lack of acceptable criteria to compare authentic assessment measures. The confusion about terminology and the lack of agreement about acceptable criteria have created doubt in the minds of some physical educators when authentic assessment data are used to make critical decisions (Lund, 1997). Yet, available literature strongly suggest that authentic assessment practices in physical education might provide opportunities that will improve students' self-concept, motivation, and skill achievement (Head, 1996; Hensley, 1997; Kirk, 1997; Kritt, 1993; Lund,

1997; Mitchell, 1992). However, there is a lack of empirical evidence to support these claims. The purpose of this study was to describe and analyze the type and the extent of authentic assessment use in public school physical education in Iowa. A second purpose was to investigate physical education teachers' perceptions about the impact of authentic assessment use on students' self-concept, motivation, and skill achievement.

Research Questions

1. What is the extent of authentic assessment use in public school physical education in Iowa?
2. What are the different types of authentic assessment techniques use in public school physical education in Iowa?
3. Are there any differences in the extent or frequency of use of authentic assessment between female and male physical education teachers at different grade levels in Iowa?
4. What is the perceived impact of authentic assessment use between female and male physical education teachers at different grade levels on student self-concept, motivation, and skill achievement.

Limitations

The following limitations were identified:

1. The study used survey technique to collect data.
2. The willingness of the subjects to complete the questionnaire.

3. The number of questionnaires returned.
4. The questionnaire was not fully validated.

Delimitations

1. The subjects were public school physical education teachers from the state of Iowa.
2. The subjects were selected for the interview.

Assumptions

The study was conducted with the following assumptions:

1. Current public school physical education teachers completed the questionnaire.
2. The individuals answering the questionnaire understood the directions and followed them appropriately.
3. The participants' responses reflected perceptions of public school physical education teachers in the state of Iowa.

Definition of Terms

For consistency of interpretation in this study, the following terms were defined:

Authentic Assessment: Assessment tasks that use multiple scoring systems to measure students' habits and repertoires on significant tasks related to life outside the classroom (Wiggins, 1989b; Worthen, 1993).

Constructivism: A theory that posits that learners construct their own knowledge through conversation, hands-on activity or experience (Ellis & Fouts, 1993; Noel, 1993).

Extent of Authentic Assessment use: Proportion of teachers who use authentic assessment.

Frequency of Authentic Assessment use: How often authentic assessment “users” employ or use authentic assessment techniques.

Motivation: The driving force or the inner-thrust behind behavior (Woolfolk, 1990).

Self-Concept: The composite of ideas, feelings, and attitudes students have about themselves (Woolfolk, 1990).

Skill Achievement: The resulting outcome of a student’s actions or technique accomplished through great effort, perseverance or courage (Dunham, 1994).

CHAPTER II

REVIEW OF LITERATURE

Have we forgotten a fact in the reform conversation? The student is the primary client of all assessment. Assessment, especially testing, should thus be designed to improve performance, not just monitor it. A preponderance of one-shot testing is never in the students' interest whether the tests are multiple-choice or performance-based. (Wiggins, 1993, p. 5)

There is a lack of empirical research about authentic assessment in physical education. The literature available is based more on opinion papers; therefore, the literature review section is based on the available opinion papers. The purpose of this study was to describe and analyze the type and the extent of authentic assessment use in Iowa public school physical education. A second purpose was to investigate physical education teachers' perceptions about the impact of authentic assessment use on students' self-concept, motivation, and skill achievement. The literature review was organized into the following discussions and subheadings:

1. Authentic assessment in general uses which included the definition of authentic assessment and the characteristics of authentic assessment.
2. Authentic assessment uses in the public schools that included the case for authentic assessment.

3. Authentic assessment uses in physical education that included the portfolio assessment, exhibitions/demonstrations, oral discourse, written essay, and the examination of the influence of the constructivist education on physical education.

Authentic Assessment in General Use

In 1983, the National Commission on Excellence in Education (NCEE) issued a comprehensive and startling report about the condition of the public elementary and secondary schools in the United States. The report concluded that the educational foundation of this country was concealed by a rising tide of mediocrity (National Commission on Excellence in Education, 1983). The nation was at a great disadvantage, especially in education and in the world marketplace. For the country's education to bounce back, as well as make the nation competitive on the global marketplace, school reform movements were asked to find better ways to assess students' work and learning in the schools (Darling-Hammond, 1994; Pullin, 1994). Thus, the calls for a paradigm shift in assessment from the traditional tests to authentic assessment (Darling-Hammond, 1994; Wiggins, 1989a; Winograd, Jones, & Perkins, 1994).

Authentic assessment is the most recent innovation in student assessment (Killoran, 1992). The benefits associated with the use of authentic tasks have made some assessment specialists (Darling-Hammond, 1990; Madaus & Kelleghan, 1993; Wiggins, 1989a) advocate

for authentic assessment use nationally. But the term authentic assessment faces definition problems. No consensus has been reached about the definition of authentic assessment. However, there is agreement on what authentic assessment is not: it is not an assessment set in a multiple-choice format (Katims et al., 1993). Because of the lack of definitional clarity, any meaningful discussion of authentic assessment without first defining the researcher's perception will cloud the issue.

Authentic Assessment Defined

Authentic assessment is a direct examination of students' performances on a significant task relevant to life outside the classroom (Worthen, 1993). Authentic assessment is a continuous feedback to students that records student learning through exhibits and work samples inherent to the school settings (Melograno, 1994).

Authentic assessment requires the presentation of exemplary tasks designed to represent performance in the field (Lund, 1997; Wiggins, 1989b). Assessment is authentic when it occurs in the context of a meaningful learning environment and reflects worthwhile learning experiences that can be documented through observation, oral presentations, anecdotal records, journals, logs, actual work samples, conferences, portfolios, discussions, experiments, projects, debates, or exhibitions along with collections of students' written products and video

tape of performances (Archbald & Newman, 1988; Lund, 1997; Parker, 1998; Winograd et al., 1994). Authentic assessments are conceptualized, complex intellectual challenges, not fragmented and static bits of tasks (Wiggins, 1989b; 1993).

Authentic assessment aims at providing students with systematic opportunities to gain ownership of and insight about their own learning, and it is intended to provide teachers with a rich basis for making professional judgment about instruction (Winograd et al., 1994).

Prevalent opinion holds that authentic assessment is nonstigmatizing, increases students' motivation, assists teachers with decision-making and is effective for reporting students' achievements and progress to parents (Melograno, 1994; 2000).

Characteristics of Authentic Assessment

Grant Wiggins promoted the concept of authentic assessment (Wiggins, 1989a). He suggested that to make a task authentic, teachers should design the performances that they wanted students to be good at before worrying about a method of grading them. Such assessment tasks require thorough thinking and planning. To make authentic assessment feasible for teachers, Wiggins (1989b; 1993) and Lund (1997) among others came out with the following seven characteristics common to all authentic tasks:

1. Authentic tasks are designed to be truly representative of performance in the field. Students should be encouraged to write, not to answer questions about writing (Darling-Hammond, 1990). Students should conduct science experiments rather than memorizing disjointed facts about science. Such conceptualized, complex intellectual challenges should involve the students' research or use of knowledge that make appropriate room for the students' learning styles, aptitudes, and interest (Wiggins, 1989b).

2. Authentic assessment processes evaluate the essentials of performance against well-clarified performance standards (Lund, 1997; Wiggins, 1989b; 1993). The performance standards are made available to the students and others in the learning community in advance rather than inferred as in the national standardized tests. The availability of the standards allows the students to know how they will be evaluated.

3. Self-assessment is an important component in authentic tasks (Wiggins, 1989b). The goal of authentic assessment is to help students to revise, modify, and take initiative to assess their own progress (Katims et al., 1993). Lund (1997) used gymnastics routines as an example of self-assessment by which students can film their performances and compare with an accepted criterion. Also, peer assessment can be done to give the students' perspectives different from that of the teacher.

4. Authentic assessment begins with the instruction and both are practically inseparable (Lund, 1997). Authentic assessment is continuous and/or formative and forms part of the instruction. The formative assessment makes the task central to students and the students are motivated to revise and improve their performances because it becomes important to them (Darling-Hammond, 1990). Such revisions lead to success as students learn from their mistakes.

5. The Latin roots of the word assessment means "to sit with" (Wiggins, 1993). The meaning of assessment makes the authentic assessment teacher an ally and not an adversary. Students work together with the teacher until mastery is achieved. The collaborations between teacher and student as well as the formative nature of authentic assessment give students the opportunity to practice and improve.

6. Authentic assessment consists of the examination of the process and product of learning (Lund, 1997). The process used in creating the product is equally important in authentic assessment because the process helps students gain ownership of their learning. In game play, for example, students should know why they choose one strategy over the other.

7. Finally, the students should present their work and defend themselves publicly and orally to ensure that they have command over their work (Lund, 1997; Wiggins, 1989b). As students present and

defend their work, they understand and appreciate the significance and importance of the work. Also, other students can learn from watching and listening to the presentation (Lund, 1997).

Authentic Assessment Uses in the Public Schools

Perhaps the characteristics proposed by Wiggins (1989b) and Lund (1997) have contributed to the widespread interest in the use of authentic assessment in the public schools particularly in the middle schools. Authentic assessment is compatible with the current middle school philosophy (Dana & Tippins, 1993). The middle school concept is built around a small community of learners where close, stable, and bilaterally respectful relations with teachers and peers are the hallmark for intellectual development and personal growth. Teaching in the middle school requires flexibility of instructional time and adequate resources in order to tap the different potentials and abilities of the students. To incorporate this flexibility into instruction requires continuous assessment that begins with the instruction (Dana & Tippins, 1993; Kritt, 1993; Lund, 1997).

In addition, the middle school students have varying abilities in growth and experience (Kritt, 1993). Students vary with rapid, uneven physical growth and development in complex skills (Dana & Tippins, 1993). The middle school teenagers appear to change rapidly over time. Motivation and persistence among these adolescents also vary. To

prepare students at this level to meet the problems and challenges that people face in real situations requires instruction and assessment that are diversified and interconnected.

The emphasis in the middle school on the interconnected nature of subjects makes the use of authentic assessment especially meaningful. To tap students' knowledge and experiences from other classes, Dana and Tippins (1993) used concept maps. Concept maps are individual's diagrammatic representation of meaningful relationships between concepts (Dana & Tippins, 1993; Novak & Gowin, 1984). In this task, students were divided into groups and asked to map their solutions to simple problems. The conclusions were that the concept map helped students use their knowledge from other classes to present ideas logically and diagrammatically (Dana & Tippins, 1993; Lund, 1997). Also, because there are no right or wrong answers in concept mapping, students feel confident to talk about their ideas expressed in the concept map.

Until recently, assessment, especially standardized testing, has existed primarily as external to the instructional process (Cole, 1992). The current changes in educational practices, instruction, and assessment expedited the need for the educational testing center to develop the "Program for Middle School Mathematics" (PACKETS; Chittenden, 1991). PACKETS is a set of classroom-based materials that

teachers use as part of their instruction. A PACKETS activity proposes a real life problem for which the solution must address the demands of a given client. Activities in this package are designed in such a way that the teacher can promote and document students' learning at the same time. When a sixth grade teacher asked the class for their views after using PACKETS, a student commented: "I think this is a good idea because using stupid problems make you wonder why this? But now I know how I can use math in the future" (Katims et al., 1993, p. 28). The above statement shows how the math activity makes sense within the student's understanding of the world. PACKETS calls for students' use of mathematical skills and knowledge in a real life context rather than presenting mathematics problems in separate bits. Also, the PACKETS materials are designed in such a way that the students can evaluate their own responses, revise their work and improve their performances.

Tinsley (1993) after many years of frustration as a high school teacher decided to teach second grade. At this grade level she realized that students are not taught about things but how to do things. Few children get it the first time when they are taught to read, count, sort, cut or paste. However, they learn from failure, watching each other and from their own first attempts. Her experience at the second grade level led to the redefinition of her views about teaching and assessment that build students' confidence in the classrooms. Her recommendations

were that good assessment should engage students in assessing their own progress. Evaluation practices must foster self-evaluation and independence with maximum feedback to students. In short, she recommended an approach to teaching and learning that demands authenticity.

The Case for Authentic Assessment

In the last two decades, the volume of traditional assessment, especially standardized multiple-choice tests, in the United States schools has increased (Neil & Medina, 1989). At the same time, complaints about traditional forms of assessment--whether it is a multiple-choice, machine-scored test or standardized sport skills and physical fitness tests--have become louder (Haney & Madaus, 1989; Hensley, 1997). The increase in the dissatisfaction has led to a search for alternate ways to measure students learning. Many acknowledged assessment authorities tout alternative assessment as the solution to the quandary (Herman et al., 1992; Wiggins, 1989b). Before the case is made for authentic assessment, let us discuss some of the common complaints about traditional standardized tests.

Haney and Madaus (1989) posited that the general case against standardized tests is that they are becoming pervasive in the schools and the same tests are used for several purposes, from diagnoses of student learning to program evaluation, and from teacher evaluation to

systemwide accountability. Traditional standardized tests, according to them, give false information about the status of learning in the nation's schools. Yet, they have been used to report students' knowledge, abilities, or skills as if these attributes can be isolated and sorted to fit on a linear scale (Darling-Hammond, 1990; Neil & Medina, 1989).

Modern theories emphasize the complexity of human intelligence. Brain-based learning research showed that the organization of the brain makes it resist isolated pieces of information unrelated to what makes sense to the learner (Caine & Caine, 1997). Any assessment of students' knowledge, abilities, or skills in isolation will produce results that are inaccurate, inconsistent, and biased (Neil & Medina, 1989).

A true test, according to Wiggins (1989b), is central to instruction, and it is made public to the students from the start because it is both central and complex. In a true assessment format, the teachers sit with their students at the end of the test to help them reason through their difficult items. Also, the teachers give the students the opportunity to retake the test if needed, which fulfills the aim of assessment as a means to improve performance, not just audit it. In addition, it makes the teacher assured that the student can accurately and willingly adapt the knowledge to varying situations (Wiggins, 1993).

On the contrary, questions on traditional standardized tests are usually kept hidden from students and teachers so that the questions

can be used again (Wiggins, 1989b). Traditional standardized tests are administered and collected without the students having a second chance to read the test questions and their answers. The students' reasons for their answers become irrelevant to the test administrators. Such mass assessment treats students as objects with similar education and thought (Wiggins, 1989b). Furthermore, it makes the assessment counter to the most basic assumptions required for learning, as students often do not have the opportunity to learn from their mistakes.

Traditional standardized tests tend to corrupt the process of teaching and learning (Haney & Madaus, 1989) and they reduce teaching to mere preparation for testing (Darling-Hammond, 1990; Wiggins, 1989b). Good teachers have students constantly moving back and forth between part and a whole performance (Wiggins, 1993), which makes students experience the feel of skillful performance and the value of developing the newer more difficult habits.

The complaints about traditional standardized assessments set the stage for the move toward authentic assessment. Authentic assessment is responsive to individual students and to school context (Wiggins, 1989b). Authentic assessment includes human judgment and discussion so that the person assessed can ask for clarification of questions and explain the answers. The authentic assessment tasks do not only reveal students' achievement to the tester, but also to the test-

taker. In addition, test takers have the opportunity to retake the test, which fulfills the aim of education, that is, to help the individual become an adept intellectual performer rather than a passive selector of traditional and prefabricated answers (Wiggins, 1993).

Authentic assessment measures do not assess students' competence on one performance but over time and in numerous performances (Wiggins, 1989b). The multiple assessments help teachers to observe patterns of success and failure and find the rationale behind them. Also, it sharpens the student's ability to prepare for and master the various roles and situations that competent professionals encounter in their work (Wiggins, 1993). Therefore, to show that a response is not an accidental response, multiple and varied assessments are required.

Another reason for the use of authentic assessment is the agreement among educators, researchers, and policymakers that traditional tests do not tap into many of the skills and abilities students need to be successful after schooling (Wiggins, 1989b; 1993). Existing traditional tests demand instruction that prepares students to memorize information and use it to solve tidy problems. Most jobs, even those requiring little skill, require technical training and flexibility (Darling-Hammond, 1990). Industries have restructured their work organizations to include cooperative planning and problem solving rather than following simple directions on the assembly line. All these changes in the

work place require education and assessment that prepare students to use resources and do complex tasks at a high level of literacy. Since the existing traditional tests get to only the basics and do not tap into higher-level literacy skills, educators, researchers, and policymakers have advocated their replacement (Dana & Tippins, 1993; Haney & Madaus, 1989; Wiggins, 1989b; 1993). Changes in society and in the work place consequently called for a move toward a new form of teaching and assessment, authentic assessment, that adequately prepares students for life (Darling-Hammond, 1990).

Authentic Assessment Uses in Physical Education

Authentic assessment in the form of performance-based assessment is not new to the physical education profession (Hensley, 1997; Melograno, 1994; Weinberg, 1996). Physical educators have used it over the years. However, the broad concern about the effectiveness of the public schools has necessitated the move toward a more authentic assessment that is linked to and is performed as an integral part of the teaching process (Hensley, 1997). In addition, there is the feeling that authentic assessment techniques may result in greater student skill achievement in physical education especially in this fast changing educational system (Wood, 1996); thus, the move toward the use of authentic assessment.

The very practical nature of physical education classes in the public schools makes the use of authentic assessment possible (Lund, 1997). However, this does not mean that every assessment conducted in the public school physical education is authentic. Assessment is said to be authentic when it uses multiple assessment scoring systems to measure students habits and repertoires (Wiggins, 1989b). Authenticity cannot be dichotomized (Hensley, 1997): it runs on a continuum with varying degrees. Consequently, some assessments in physical education are considered more authentic than the others.

Lund (1997) used two swimming pool lifeguards' preparation to illustrate the continuum of authenticity in physical education assessment. One of the lifeguard's training comprised classroom work, readings, passing a written test, and practicing a life saving situation. The other lifeguard's training has all the requirements of the former lifeguard's except the practice section. At the end of this scenario, she asked which of the lifeguards a person would select if he or she happens to use the pool. Obviously, both lifeguards have some training. However, it makes sense to select the one with additional training in the practice to rescue a victim because aside from all the knowledge that the lifeguard might have, the rescue has to be done practically. The lifeguard illustration shows the range of authenticity in assessment that exists in the public school physical education classrooms.

In physical education, it is common to see teachers assess students' skills with performance-based tests (Weinberg, 1996). Some assess students at the end of a class unit whereas others assess at the beginning and end of the course. For example, a volleyball class might assess students' skills on the bump, the set, and the service. A basketball class might time students dribbling balls through cones. However, in the test formats described above, the contexts in which the assessment is conducted and the task performed are artificial in nature (Hensley, 1997). The assessment context bears little resemblance to how the skill will be used in the real sport context. Because they are decontextualized and different from real game situation, applying them efficiently in a real game situation becomes a problem (Weinberg, 1996). As much as skills are needed for effective game performance, the situation for learning and assessment should be identical to the actual game situation to make it more authentic and easily transferable.

The field of physical education provides several opportunities for students to demonstrate behaviors that are directly observable and measurable. The performance of a psychomotor skill, for example a tennis serve, makes observational analysis and subjective judgment a logical choice for physical educators (Hensley, 1997). Imwold, Rider, and Johnson (1982) surveyed the methods used to determine and report students' performance in the Florida public school physical education

programs. Of the 500 physical education teachers they surveyed, 29% of the males and 50.5% of the females used authentic assessment. A similar survey by Hensley et al. (1989) revealed that only about 40% of the physical education teachers used standardized tests. Therefore, it can be inferred that the remaining 60% used alternative assessment, which includes authentic assessment and teacher made tests. Another investigation of measurement practices confirmed that physical educators relied more on subjective assessment practices (Hensley, 1997). The reliance on subjective assessment practices indicated that physical educators use assessment other than the traditional standardized tests regardless of the criticism against alternative assessments on the basis of validity and reliability. Such findings in addition to the 1995 National Physical Education Standard (NASPE) publication, *Moving into the Future*, have precipitated interest in the use of authentic assessment among public school physical educators. At present, Lund (1997) and Parker (1998) have proposed the use of portfolios, oral discourses, exhibitions/demonstrations, written essays, checklist, event tasks, student logs, peer observation, and parental report as some of the authentic assessment techniques that could be used in physical education. The next section of the review will focus on some authentic assessment techniques used in physical education.

Portfolio

Professionals in the creative arts have used portfolios to demonstrate the success of their works for many years (Hauge, 1997). Recently, educators have recognized portfolios as an important means through which students can demonstrate their success and academic progress (Dana & Tippins, 1993; Hauge, 1997). A portfolio is a collection of students' work over time that exhibits their efforts, progress and achievements (Dana & Tippins, 1993; Feuer & Fulton, 1993; Hauge, 1997; Herman & Winters, 1994; Kirk, 1997; Lund, 1997; Melograno, 1994; 2000; Meyer & Schuman, 1990). Portfolios provide a more equitable and sensitive portrait of what students know and are able to do. Portfolios can become a window into the students' heads, a means through which teachers, parents and students can understand the educational process at the level of the individual (Paulson, Paulson, & Meyer, 1991).

Portfolios are valuable ways to link instruction and assessment (Kirk, 1997), and can serve valuable purposes in physical education. With portfolios, students prepare and select their best works that should be graded and included in the portfolio (Melograno, 2000). Because the portfolio is a continuous collection of students' work, the students are actively involved in the learning process.

In 1995, NASPE, recommended portfolio use in physical education (NASPE, 1995). The four areas for portfolio use cited by NASPE were students' growth, student thinking, students' views of themselves as learners, and problem solving. Today, most physical educators will agree that portfolio use does not only increase student self-concept and motivation, but accommodates individual differences and encourages critical thinking (Hauge, 1997; Lund, 1997).

Also, physical educators can use portfolios to document students' fitness gains (Lund, 1997). Students can use the portfolio to show knowledge of their fitness gains components and develop an activity program that will achieve a fitness goal. Logs can be used to record individual workouts with pre - and post-testing documenting fitness improvement. The use of portfolios encourages individual learning, documents achievement outcomes, provides feedback and continuous evaluation of student progress, and can showcase physical education program and student achievement by displaying students' work during various school events (Kirk, 1997; Lund, 1997). However, despite NASPE's recommendation for the use of portfolios and the benefits that professionals in other fields have achieved with portfolios, the literature does not reveal much support for portfolio use in physical education.

Exhibitions/Demonstrations

Exhibitions and demonstrations are broad presentations of students' performance (Feuer & Fulton, 1993; Lund, 1997). Exhibitions are one of the authentic assessment formats that incorporate all aspects of the physical education process - - physical, mental, emotional, affective, and social (Bloom et al., 1956). Let us consider, for example, a novice student who wants to tumble in gymnastics. Before the student can demonstrate, the teacher helps plan the various phases of the skill. The student becomes aware of the acceptable criterion of performance as the teacher walks the student through the many phases of the skill. And, as the student learns the skill, feedback is received from the teacher. A videotape of the student's performance can be played back for analysis and correction. The active role the student plays in the planning, setting up the equipment, and the time invested in the activity before mastery is achieved, makes the learning of the skill rewarding. Also, because the student has the opportunity to move forward and backward in every phase of the skill, there is not much pressure to perform. Furthermore, the assessment and evaluation of the skill are done in conjunction with the student, which does not only motivate or increase skill level and/or self-confidence of the students (Kirk, 1997; Lund, 1997; Zhu, 1997), but makes them gain control and ownership of their learning. Such an approach to instruction and assessment is

congruent to the performance expected of every individual in the work place: that is, planning, execution, rehearsal/editing, feedback from allies, and mastery (Wiggins, 1989b).

Oral Discourse

Another form of authentic assessment use in physical education is oral discourse. Oral discourse gives students the opportunity to demonstrate their competence and knowledge (Lund, 1997). Oral discourse requires students to develop their knowledge, synthesize their thoughts, present their ideas, draw conclusions, and defend their choices. Through oral discourse, for example, students can discuss the merits and demerits of zone versus player marking defense. When students are given the opportunity to discuss their knowledge in class, they not only improve their rhetorical abilities and self-confidence, but also become critical listeners and consumers of information. Through oral discourse, students can check their knowledge against others.

Written Essay

Essays are used to describe and explain facts in context. To present the material coherently requires analysis, synthesis, and critical thinking skills. Essays are not limited to traditional papers; they can be used in physical education, too. Lund (1997) talked about a student teacher who used essays to help the class record facts about the heart during a heart awareness month. The teacher created a heart obstacle

course that emphasized fitness and heart facts acquired through physical activity. As the weeks passed, the teacher made the obstacle course difficult by including additional heart facts and physical activities. The students' tasks were to record these facts, research and read some literature about the heart and create a brochure with the information gathered. The brochures were distributed to parents at an open house. The conclusions were that the students integrated their knowledge in language arts, technology, and physical education to talk about the heart in relation to physical activity to the parents. When the students displayed their work during a parent conference, their work informed the community about the content of physical education and wellness (Lund, 1997). The opportunity the students had in creating the brochure about the heart made them become owners of their knowledge. For a teacher to guide students to accomplish these performances requires a teaching/learning classroom that is learner-centered and constructivist in nature. The next section of the literature review will consider the constructivist theory and how it has helped physical education teachers reconsider the use of authentic assessment in the classroom.

Constructivism

Many cognitive models and approaches to teaching and practice are presented in the literature. One such cognitive approach that is growing in popularity is constructivism. Constructivists believe that

each learner approaches a learning task with a different set of personal beliefs, motivations, and conceptions (Holmes & Leitzel, 1993) and those learners actively construct their own knowledge (Bodner, 1986). Because of the active involvement, learners compare what they are taught with their existing conceptions and frameworks of knowledge.

Constructivism emerged from the works of Piaget and Vygotsky (Manus, 1996; Vygotsky, 1986). Vygotsky (1986) maintained that thought is derived from both the experiences and maturation processes of an individual. A person's consciousness emerged from mediated activities that would then be internalized into prominent forms of cognitive functioning. In much the same way an individual's experiences partially determine a person's cognitive development. Furthermore, the social interaction between the teacher and the student form the foundation of the learner's cognitive development. In addition, it helps the student progress from understanding the basic skills to greater complexity through the guided participation of the teacher. The teacher and the student achieve this goal through intersubjectivity.

According to Piaget's stages of physical development, most children acquire basic movement patterns before reaching school going age (Woolfolk, 1990). By the age of six, the average child can walk, run, jump, skip and perform other basic movement patterns with little difficulty. As children attend physical education classes, it is imperative

that the physical education teacher organizes the classroom environment to help the learners develop their basic movement patterns.

Accomplishing this task requires teaching and learning that centers on the learners. Such an approach is congruent with the constructivist viewpoint about teaching and learning.

Other researchers, including Pankratius and Young (1995), suggested that learners use previous knowledge to construct associations that are refined continuously over time. Cobb and Steffe (1983), investigating the phenomenon of learning mathematics, concluded that the learning of mathematics involves both the teacher's and the student's efforts to make meaning of verbal and nonverbal actions or activities. Therefore, teachers should be concerned with what it takes for students to know subject matter as well as what it takes to foster understanding (Prawat, 1992).

There is also no particular way to learn in the constructivist classroom (Winn, 1991). Students use inherent, creative capabilities to determine what to learn and how to learn it. Such approach to learning enhances the acquisition of advanced knowledge. To implement this approach to learning in the classroom, according to Winn, requires instructional designers to abandon the mechanical application of instructional design procedures.

Learning is an active, continuous process that requires value judgments from the learner. Constructivist education places the learner at the center of the learning process with the teacher as a facilitator. The teacher creates and enriches the learning environment using a variety of enjoyable activities. The teacher's facilitating role helps the student to assume ownership and responsibility for learning (Alcove & McCarty, 1992).

Brooks and Brooks (1993) outlined some of the characteristics common in the constructivist classroom. In the constructivist classroom, students are encouraged to be autonomous and to take initiative in hypothesis development and questioning (DeVries, 1987). Also, relationships and metaphors are developed. Students' understanding of concepts takes precedence over teachers' views and the learning cycle is frequently used to develop curiosity. Constructivist teachers emphasize individual choice, multiple methods, and alternative approaches, all of which are better measured with authentic assessment techniques.

According to Wheatley (1991), two main principles drive the constructivist theory in the classroom. The first principle states that knowledge development evolves through active participation. Meaning is built through active involvement as the individual interacts with the content and the environment. Educators cannot put ideas into learners' minds. Instead they should create environments and opportunities that

enable the learner to evoke meaning. The first principle is fairly accepted in constructivism. The second principle by Wheatley (1991) suggested that no collective reality exist; that is, truth cannot be found. As a result, one's picture of the world is shaped by one's personal experience. The second principle is less accepted in constructivism.

Applying Constructivism in Physical Education

Taking Wheatley's (1991) definition of constructivism and applying it to physical education, one can see how the constructivist approach fits well with physical education and the practice of authentic assessment. Physical education in the K-12 school system is very practical and process oriented. The practical and process natures offer opportunities for the student to exhibit behaviors at the student's level of development that may be directly observed by the teacher. For example, the assessment of a motor task in the psychomotor domain is mostly done through the use of observational analysis and subjective judgment (Hensley, 1997). Such observational analysis provides an opportunity for the teacher to use multiple assessment techniques to evaluate the student performance.

Many physical educators use different assessment techniques to evaluate their students' performances as the students learn through active participation. The common practice used in most K-12 physical education skill assessment is the pretest-posttest assessment format,

whereby a student is assessed at the beginning and at the end of the course unit. For example, Hensley (1997) developed a scoring rubric for the forehand and backhand drive in tennis. According to him using such rubrics in physical education measures discrete skills, which are for diagnostic purposes. His position was that rather than assessing isolated skills, teachers should use assessment techniques that touch on both the holistic and discrete skills. Adopting Hensley's (1997) position toward physical education assessment ties well with Wheatley's (1991) constructivist principle whereby knowledge is developed through active participation. To accomplish this in physical education requires the application of the constructivist philosophy of teaching and learning where students are guided to construct their own knowledge with ample opportunities to demonstrate their abilities. To achieve this requires a shift in teaching and assessment from the traditional approaches to authentic assessments that use multiple methods and alternative approaches to learning and assessment.

Summary

Constructivist theorists postulate that learners develop knowledge through active participation. Learning to the constructivist is an active continuous process that requires value judgments from the learner. Based on the constructivist theory assessment should help learners improve and not just monitor performance.

Although available literature indicates that authentic assessment use will improve students' self-concept, motivation, and skill achievement in physical education, there is no empirical data in the literature to support these claims. This study is intended to fill the gap in the literature.

CHAPTER III

METHODS

The purpose of this study was to describe and analyze the type and the extent of authentic assessment use in Iowa public school physical education. In addition, a second purpose was to investigate physical education teachers' perceptions about the impact of authentic assessment use on students' self-concept, motivation, and skill achievement.

Design

The research design for this study was descriptive. A mail-out questionnaire and a face-to-face interview approach were used to collect data. The survey method of data collection was considered appropriate for the study because of the need to gather large amounts of data from respondents in their natural settings (Wimmer & Dominick, 1997). In addition, the widespread location of the 1,934 public school physical education teachers within the state of Iowa makes it difficult to gather data from the physical education teachers without the use of the mail survey technique.

The face-to-face interviews helped the researcher understand the participants' perspectives from their practical points of view (Glesne & Peshkin, 1992). In addition, the interview data helped confirmed the survey findings. The questionnaire and the interview techniques were

appropriate to help the researcher understand the physical education teachers' perceptions about the use of authentic assessment.

Subjects

Sampling

Three hundred and ninety-six physical education teachers in the state of Iowa were randomly sampled. The teachers were classified as either elementary or middle/junior high or high school physical educators. A stratified random sampling technique was used to select 132 female and male teachers from each grade level. The stratified random sampling technique ensured equal representation of subjects by gender and by grade level (Huck & Cormier, 1996; Krathwohl, 1993). In addition, the equal representation of subjects was designed to minimize potential standard error of the difference (Rossi, Wright, & Anderson, 1989), and for accurate comparisons of subjects by gender and grade level.

Demographic Information

The subjects answered a demographic questionnaire about their gender and educational level, the current grade level teaching, the total number of years taught, the number of years teaching physical education, and the number of years in present school. For mailing purposes, the names, the school addresses and grade level of subjects

were obtained from the Iowa Department of Education in Des Moines, Iowa.

Dependent Measure

Physical Education Authentic Assessment Inventory (PEAAI)

A questionnaire, Physical Education Authentic Assessment Inventory (PEAAI), was used to collect data for the mail-out survey part of the study. The questionnaire contained two sections. Section A of the questionnaire requested the subjects to rate the extent to which they use each of fifteen authentic assessment techniques in their physical education classes. The respondents' ratings were from always use (A), often use (O), sometimes use (S), and rarely use (R) to never use (N).

Section B of the questionnaire requested the physical education teachers to rate their perceptions about the impact of authentic assessment use on their students' self-concept, motivation, and skill achievement. The Section B of the questionnaire had three subscales. The subscale that measured perceptions of effect on self-concept comprised items 3, 4, 8, 12, and 13. The subscale items 2, 6, 7, 11, and 15 measured perceptions of effect on motivation, and subscale items 1, 5, 9, 10, and 14. The directions on the Section B of the PEAAI questionnaire requested subjects to respond with answers of strong agreement (SA), agreement (A), disagreement (D), and strong disagreement (SD). The Section B of the questionnaire yielded three

separate scores: one each for self-concept, motivation, and skill achievement.

In order to establish internal consistency reliability for the PEAAI, Cronbach Alpha was computed for each of the three subscales. For the subjects in this study, alpha coefficients were .55 for motivation, .64 for skill achievement, and .74 for self-concept, indicating moderate levels of internal consistency.

Pilot Testing

Two pilot tests (participating pilot test and undeclared pilot test) were conducted prior to formal data collection. Although there is no specific order to follow when pilot testing a questionnaire, Converse and Presser (1986) recommended that the participating pilot test precede the undeclared pilot test. The two pilot tests are discussed in the sections that follow.

Section I: Participating Pilot Test

The participating pilot test was a discussion of the questionnaire with the participants. Four physical educators participated in the participating pilot test. Three of the subjects teach physical education to undergraduate and graduate physical education students in a medium size Midwestern University. The fourth subject was the Head of a Physical Education Department in a high school.

In the participating pilot test, the subjects were told that it was a practice run and that they should feel free to comment on the questionnaire items. Separate interviews were conducted with each subject to discuss each one's comments, reactions and responses. Such a subjective, organized review of the questionnaire's contents by the subjects ensured that the questionnaire included relevant content areas of interest (Converse & Presser, 1986). In addition, the participating pilot test helped the researcher identify and delete duplicate items, and allowed for clarification of those items that were confusing to the subjects.

Section II: Undeclared Pilot Test

The undeclared pilot test was conducted in April 11 and 12, 1999, at the 11th Annual Physical Education PK-12 Learning/ Sharing Conference in the Wellness/ Recreation Center, University of Northern Iowa in Cedar Falls, Iowa. Seventy practicing physical educators who attended the conference were given copies of the questionnaire. Thirty-five (50%) of the participants completed and returned the questionnaire.

Permission was obtained from the conference organizer for the use of the conference participants as subjects for the pilot test. The PEAAI questionnaire and a short introductory letter were handed to subjects by the researcher at the registration desk as conference participants signed in. The introductory letter attached to the questionnaire stated the

purpose of the study, the approximate time it would take each participant to complete the questionnaire and the instructions for questionnaire return.

Procedure

Application for permission to conduct the study was filed with the Institutional Review Board of the University of Northern Iowa prior to data collection. The application form included the title of the research, the name of the principal investigator, explanation of the test, and the risks and discomfort subjects might encounter for participating in the study. Approval to conduct the study was received prior to data collection.

Data Gathering

Data for the study were gathered in two phases. Phase I was the mail survey. Phase II used face-to-face interviews with selected subjects.

Phase I: Mail Survey

A cover letter (Appendix A), teacher demographic questionnaire (Appendix B), the PEAAI questionnaire (Appendix C), and a self-addressed prepaid envelope were mailed to the selected physical education teachers. The letter requested the voluntary participation of the physical education teacher in answering the questionnaire. The letter briefly described the purpose of the study and assured complete confidentiality of the responses. Included in the cover letter was an

estimate of the time needed to complete the questionnaire. No information that could identify the individual teachers was requested.

Phase II: Interview

A focus interview approach was used to gather data from the interview participants. The focus interview was a combination of the exploratory and structured interview techniques that allowed the respondents to set the initial course of the interview and increasingly focus on the researcher's agenda as the interview progressed (Krathwohl, 1993).

Twelve physical education teachers (4 elementary school, 4 middle/junior high school, and 4 high school) who participated in the mail survey were selected for the interview. In order to ensure an equal representation of subjects by gender at each grade level, a purposive sampling technique was used to select the physical education teachers for this second phase of the study (Krathwohl, 1993; Rea & Parker, 1992).

The selected physical education teachers were contacted for their permission for the interview. Each physical education teacher who agreed to participate in the interview signed an informed consent form (see Appendix D). Field notes and a hand cassette recorder were the means for data gathering during the interview process. Each physical education teacher was interviewed once. The interview was used to

examine teachers' perceptions about the impact of authentic assessment use on students' self-concept, motivation, and skill achievement. The interview questions included, but were not limited to the following:

1. How do you assess your students? Please give me examples.
2. What do you call these types of assessment? For example, authentic or performance or alternative assessment.
3. How long have you used this form of assessment?
4. Do you also use traditional assessment measures - paper and pencil tests?
5. Why do you use authentic or performance or alternative assessment measures?
6. Some advocates think authentic assessment has increased students' self-concept, motivation, and skill achievement. Do you think using authentic assessment has improved your students' self-concept, motivation, and skill achievement? Please give examples.
7. What assessment technique(s) did you use before using authentic assessment? Why did you bring authentic assessment measure into your assessment approach?

Description of Data

Scoring

For the purpose of scoring and data analysis, each assessment technique of the section A of the PEAAI was scored separately. Each

assessment technique was given numerical values indicating level or degree of use as follows: always use (5 points), often use (4 points), sometimes use (3 points), rarely use (2 points), and never use (1 point).

The responses to the questions in the section B of the PEAAI were assigned a four-point Likert type scale for the purpose of scoring and data analysis. Each question had a score order of strongly agree (5 points), agree (4 points), disagree (3 points), and strongly disagree (2 points) except for items 3, 5, 9, 11, 13, and 15 which had a reverse score order of strongly agree (2 points), agree (3 points), disagree (4 points), and strongly disagree (5 points). Because there are an equal number of items in each subscale, the subscale score is obtained by summing the ratings in each scale. Thus, the range for each subscale was from a high of 25 points to a low of 10 points. The final scores for each subscale reflected the physical education teachers' perceived impact about authentic assessment use on students' self-concept, motivation, and skill achievement.

Data Analysis

Mail Survey

Descriptive statistics were calculated on the respondents' demographic information. In order to determine the extent of authentic assessment use among physical education teachers, frequency statistics were calculated. Additional statistics were calculated to describe the

types of authentic assessment use, and the frequency with which these types were used. Separate Kruskal-Wallis tests was calculated on the reported level of use of each popular authentic assessment technique to investigate differences in usage among grade levels. Additional statistics were calculated to investigate differences on the reported level of use of each popular assessment technique by gender.

Three separate One-Sample t-tests were calculated on the physical education teachers' perceived impact scores to assess the nature (positive or negative or neutral) of the impact. A 2 x 3 Multivariate Analysis of Variance (MANOVA) was calculated to investigate differences in the respondents' perceived impact scores on students' self-concept, motivation, and skill achievement.

Data were analyzed using the Microsoft Windows 95 PC version of the Statistical Package of the Social Sciences (SPSS). An alpha of $p < .05$ was set as the level of significance for each statistical test. Because this study was descriptive and not many significant differences were found the likelihood of an inflated alpha level was not a major concern.

The interview data analysis began with the first data collected and was an ongoing process. The data were analyzed based on the content and emerging themes as recommended by Lincoln and Guba (1985).

CHAPTER IV

RESULTS

Although many research methods text including Krathwohl (1993) recommend about 80% return rates in questionnaire studies, the 53.0% return rate was substantial for this study. The result section of this study was based on the 53.0% who returned the questionnaire.

The purpose of this study was to describe and analyze the type and the extent of authentic assessment use in Iowa public school physical education. In addition, another purpose was to investigate physical education teachers' perceptions about the impact of authentic assessment use on students' and self-concept, motivation, and skill achievement.

Subject Demographics

Out of the 396 physical education teachers surveyed, 210 (53.0%) questionnaires were returned. One hundred and two (48.6%) of the subjects who returned the questionnaires were females and 108 (51.4%) were males. Eighty (38.1%) of the subjects taught at the elementary school level, 70 (33.3%) taught at the middle/junior high school level, and 60 (28.6%) taught at the high school level. The subjects' teaching experience ranged from 1 year to 35 years (\underline{M} = 18.12, \underline{Mdn} = 20.0). Years of physical education teaching ranged from 1 year to 35 years (\underline{M} = 16.24, \underline{Mdn} = 17.0). The teachers' years of authentic assessment use

ranged from 1 year to more than 10 years. The teachers' educational level ranged from the bachelor to the doctorate degree levels.

Extent of Authentic Assessment Use

For the purpose of this study, extent of authentic assessment use is defined as the proportion of teachers who use authentic assessment. In order to determine the extent to which physical education teachers use authentic assessment, frequency statistics were calculated. Results indicated that 158 (75.2%) teachers involved in this study use authentic assessment. The teachers who use authentic assessment included 82 (51.9%) females and 76 (48.1%) males. There were 71 (44.9%) elementary school teachers, 47 (29.7%) middle/junior high school teachers, and 40 (25.3%) high school teachers.

The remaining 52 (24.8%) physical education teachers indicated that they did not use authentic assessment techniques. The teachers who do not use authentic assessment were asked to give their rationale for not using authentic assessments, but asked no further questions.

The teachers' rationales for not using authentic assessment vary. For some, (28.8%) authentic assessment techniques consume time. In addition, as one respondent stated, "because my physical education classes meet once every week, it is difficult to incorporate authentic assessment practices into my program." Others (30.8%) indicated that the number of physical education teachers in their programs had been

reduced. Furthermore, some (19.2%) teachers indicated they teach every hour of the school day and supervise students during lunch. Taken together the reduction in the number of teachers and/or the big class sizes, it is not surprising that the teachers find it difficult to incorporate authentic assessment practices in their programs. For the rest (21.2%), authentic assessment techniques were reportedly very new and they did not have the opportunity to use them in their teachers' preparation programs. Thus, they felt uneasy in using authentic assessment.

Reported Use of Individual Authentic Assessment Techniques Among
"Users" by Gender

In this study, "frequency of authentic assessment use" is how often authentic assessment "users" employ authentic assessment techniques. In order to determine the frequency with which the physical education teachers used each type of authentic assessment technique, the teachers were asked to rate each of fifteen authentic assessments on a scale (1 = never use through 5 = always use). Each physical education teacher who indicated any use of any authentic assessment technique was considered a "user" of the technique and was included in the subsequent analysis.

For the female participants, 100% of the teachers reported using teacher observation (see Table 1). In addition, a high percentage of the female teachers reported use of self-observation (97.6%), checklist (95.1%), peer observation (93.9%), demonstration (91.5%), and event

task (86.6%) techniques. Taken together, all of the male participants reported using teacher observation (100%). In addition, a large percentage of the males reported using self-observation (93.4%), checklist (92.1%), demonstration (89.5%), event task (86.8%), and peer observation (85.5%) techniques.

Overall, a limited number of female and male physical education teachers reported using the essay and the portfolio. For the female teachers, 41.5% and 43.9% reported they use the essay and the portfolio, whereas 44.7% and 42.1% of the males also reported they use the essay and the portfolio, respectively (see Table 1).

A further examination of the data in Table 1 suggested differences in the proportion of authentic assessment use by gender. However, tests of comparison on the proportions of use did not show any significant statistical difference in female and male usage of any of the authentic assessment techniques, $\chi^2(1) = .23, p > .05$. Overall, female and male physical education teachers did not differ significantly on the proportion of authentic assessment use (see Table 1).

Table 1

Reported Use of Individual Authentic Assessment Techniques Among
“Users” by Gender

Assessment	Total (<u>N</u> = 158)	Female (<u>n</u> = 82)	Male (<u>n</u> = 76)
Teacher Observation	100% (158)	100% (82)	100% (76)
Self Observation	95.6% (151)	97.6% (80)	93.4% (71)
Checklists	93.7% (148)	95.1% (78)	92.1% (70)
Demonstration	90.5% (143)	91.5% (75)	89.5% (68)
Peer Observation	89.9% (142)	93.9% (77)	85.5% (65)
Event Task	86.7% (137)	86.6% (71)	86.6% (66)
Group Project	76.6% (121)	78.0% (64)	75.0% (57)
Student Project	74.7% (118)	79.3% (65)	69.7% (53)
Oral Discourse	74.1% (117)	70.7% (58)	77.6% (59)
Parental Report	71.5% (113)	70.7% (58)	72.4% (55)
Video	68.4% (108)	73.2% (60)	63.2% (48)
Student Log	55.1% (87)	58.5% (48)	51.3% (39)
Anecdotal Record	54.4% (86)	50.0% (41)	59.2% (45)
Portfolio	43.0% (68)	43.9% (36)	42.1% (32)
Essay	43.0% (68)	41.5% (34)	44.7% (34)

Note. No significant difference was found, $p > .05$.

Reported Use of Individual Authentic Assessment Techniques Among
“Users” by Grade Level

Extent of authentic assessment use was calculated on each of the 15 assessment techniques by grade level (see Table 2 for data). The data in Table 2 showed that 100% of the participants who use authentic assessment reported using teacher observation. In addition, large percentages of the teachers, across all grade levels, reported using self-observation (95.6%), checklists (93.7%), demonstration (90.5%), peer observation (89.9%), and event tasks (86.7%). Overall, only a small number of the teachers, at all grade levels, reported using the essay (43.0%) and the portfolio (43.0%) techniques. Thus, most physical education teachers' use all those authentic assessment techniques that are patently appropriate for overt physical and in “real game” performance.

Further examination of the data in Table 2 showed few statistically significant differences in the proportions of authentic assessment use across grade levels. A chi-square test of comparison of the proportion of usage showed a significant difference between the elementary and the high school physical education teachers, $\chi^2 (1) = 3.87, p < .05$. Therefore, a higher proportion of elementary physical education teachers uses checklists than the high school teachers. No significant difference was found among grade levels for any of the remaining assessment.

Table 2
Reported Use of Individual Authentic Assessment Techniques Among
“Users’ by Grade Level

Assessment	Elm. (<u>n</u> = 71)	Mid/Jr. (<u>n</u> = 47)	High. (<u>n</u> = 40)
Teacher Observation	100% (71)	100% (47)	100% (40)
Self Observation	98.6% (70)	93.6% (44)	92.5% (37)
Checklists	98.6% (70) ^a	91.5% (43)	87.5% (35) ^a
Demonstration	93.0% (66)	91.5% (43)	85.0% (34)
Peer Observation	94.4% (67)	83.0% (39)	90.0% (36)
Event Task	85.9% (61)	87.2% (41)	87.5% (35)
Group Project	77.5% (55)	72.3% (34)	80.0% (32)
Student Project	71.8% (51)	72.3% (34)	82.5% (33)
Oral Discourse	76.1% (54)	78.7% (37)	65.0% (26)
Parental Report	74.6% (53)	72.3% (34)	65.0% (26)
Video	63.4% (45)	74.5% (35)	70.0% (28)
Student Log	53.5% (38)	55.3% (26)	57.5% (23)
Anecdotal Record	59.2% (42)	51.1% (24)	50.0% (20)
Portfolio	40.8% (29)	36.2% (17)	55.0% (22)
Essay	33.8% (24)	51.1% (24)	50.0% (20)

Note. Percentages in the same row that share superscripts differ at $p < .05$.

Reported Frequency of Use of Popular Authentic Assessment Techniques
by Grade Level

The data in Tables 1 and 2 showed teacher observation, self-observation, checklist, demonstration, peer observation, and event tasks as the authentic assessment techniques used by most physical education teachers. Given that most physical education teachers use the above six assessment techniques, the next section of the analysis investigated differences in the frequency with which female and male teachers at different grade levels employed the popular authentic assessment techniques.

In order to investigate differences with which physical education teachers on average employed the most popular authentic assessment techniques, subjects' responses were sorted into use (coded = 2 through 5) and never use (coded = 1). Kruskal-Wallis tests were then conducted to investigate the differences. The reported level of use of each of the popular six authentic assessment techniques was the dependent measure. Subjects indicated use on a scale of 5 = always, 4 = often, 3 = sometimes, 2 = rarely, 1 = never. Table 3 contains data for the popular authentic assessment techniques.

Table 3

Means and Standard Deviations of Frequency of use of the Popular
Authentic Assessment Techniques by Grade Level

Technique		Elm. (<u>n</u> = 71)	Middle/Jr. (<u>n</u> = 47)	High. (<u>n</u> = 40)	Total (<u>N</u> = 158)
Teacher observation	<u>M</u>	4.59	4.68	4.45	4.58
	<u>SD</u>	0.66	0.47	0.74	0.64
	<u>n</u>	71	47	40	158
Self observation	<u>M</u>	3.25	3.09	3.27	3.21
	<u>SD</u>	0.81	0.83	0.69	0.78
	<u>n</u>	70	44	37	151
Demonstration	<u>M</u>	3.59	3.55	3.41	3.53
	<u>SD</u>	0.85	0.90	0.89	0.87
	<u>n</u>	66	43	34	143
Peer observation	<u>M</u>	3.02	2.87	2.94	2.96
	<u>SD</u>	0.62	0.69	0.67	0.65
	<u>n</u>	67	39	36	142
Event task	<u>M</u>	3.24	3.31	3.48	3.32
	<u>SD</u>	0.80	0.84	0.91	0.84
	<u>n</u>	61	41	35	137
Checklists	<u>M</u>	3.21 ^a	3.20	3.68 ^a	3.32
	<u>SD</u>	0.91	0.86	0.79	0.89
	<u>n</u>	70	43	35	148

Note. Means in the same row that share superscripts differ at $p < .05$.

Teacher Observation

The results of the Kruskal-Wallis nonparametric analysis of variance on teacher observation indicated no significant differences

among grade levels, $H(2) = 1.85$, $p > .05$. The total mean for teacher observation as shown in Table 3 was high, ($\underline{M} = 4.58$). On the scale used, this value falls between always use (5) and often uses (4). Overall, the physical education teachers indicated they use teacher observation very often.

Self Observation

The results of the Kruskal-Wallis test on self-observation showed no significant differences among grade levels, $H(2) = 1.74$, $p > .05$. The total mean rating for self-observation as shown in Table 3 was moderate, ($\underline{M} = 3.21$). Comparing the mean with the scale used, this value falls on the point 3 = sometimes use. Thus, the physical education teachers sometimes use self-observation.

Demonstration

The results of the Kruskal-Wallis test on demonstration and grade level showed no significant difference, $H(2) = .93$, $p > .05$). The total mean for demonstration as shown in Table 3 was moderate, ($\underline{M} = 3.53$). On the scale used, this value falls between the scale points 3 = sometimes use and 4 = often use. Therefore, the physical education teachers reported they use demonstration quite often.

Peer Observation

The results of the Kruskal-Wallis nonparametric analysis of variance on peer observation and grade level showed no significant

variance on peer observation and grade level showed no significant difference, $H(2) = 1.79$, $p > .05$. The total mean for peer observation was low, ($M = 2.96$) (see Table 3). On the scale used, this value falls on the points between rarely use (2) and sometimes uses (3). On the average, the physical education teachers use peer observation sometimes.

Event Task

The results of the Kruskal-Wallis analysis indicated no significant differences among grade level on the use of event task, $H(2) = 1.40$, $p > .05$. The total mean for event task as shown in Table 3 was moderate, ($M = 3.32$). On the scale used, this value falls between the points 3 = sometimes use and 4 = often uses. Therefore, the physical education teachers indicated they use event task sometimes.

Checklists

The results of the Kruskal-Wallis H test on checklists and grade level was significant, $H(2) = 7.82$, $p < .05$. A Mann-Whitney U post-hoc follow-up test comparing the mean ranks showed that high school teachers ($M = 63.50$) in this study used checklists more frequently than the elementary school teachers ($M = 47.50$). The overall mean rating for checklists as shown in Table 3 was moderate, ($M = 3.32$). On the scale used, the mean value falls between the points 3 = sometimes uses and 4 = often uses. Thus, the physical education teachers sometimes use checklists.

Reported Frequency of Use of Popular Authentic Assessment Techniques
by Gender

Further statistical analyses were also conducted to investigate differences between the genders in the frequency with which each of the popular authentic assessment techniques was used. Results of the Mann-Whitney *U* analyses did not show any significant differences between gender in the frequency of use of the popular authentic assessment; all *p* values were > .05 (see Table 4 for data).

Table 4

Means and Standard Deviations of Frequency of use of the Popular
Authentic Assessment Techniques by Gender

Technique	Gender			
	Female (<i>n</i> = 82)		Male (<i>n</i> = 76)	
	M	SD	M	SD
Teacher observation	4.62	0.53	4.53	0.73
Self observation	3.30	0.78	3.11	0.78
Demonstration	3.50	0.90	3.57	0.85
Peer observation	3.02	0.68	2.89	0.61
Event task	3.21	0.87	3.45	0.80
Checkilists	3.34	0.88	3.30	0.90

Note. No significant difference was found, *p* > .05.

Perceived Impact of Authentic Assessment Use on Students' Self-
Concept, Motivation, and Skill Achievement

The present study also sought to find out whether physical education teachers perceived authentic assessment to have an impact on students' self-concept, motivation, and skill achievement. In the section B of the Physical Education Authentic Assessment Inventory (PEAAI), the respondents indicated their perceptions about the impact of authentic assessment use on a Likert scale of 5 = strongly agree, 4 = agree, 3 = disagree, and 2 = strongly disagree. The section B of the Physical Education Authentic Assessment Inventory (PEAAI) consisted of three subscales of five items each. Because there are equal numbers of items in each subscale, the subscale score is obtained by summing the ratings in each scale. Thus, the range for each subscale is from a high of 25 points to a low of 10 points. The midpoint or the neutral score was 17.5. The neutral score is obtained by multiplying the midpoint value of 3.5 by 5, the number of items in each scale.

In order to assess whether the physical education teachers perceived authentic assessment use has made a positive or negative or no impacts on students' self-concept, motivation, and skill achievement, three separate one-sample t-tests were conducted. The physical education teachers' ratings of the perceived impact were the dependent

measures. The mean scores on each of the three subscales were then compared with the test value of 17.5, the neutral score, to determine the impact.

Self-Concept

The result of the one-sample t-test on perceived impact on self-concept was significant, $t(157) = 9.26, p < .05$. The magnitude of the difference was large, $d = 0.74$ (Cohen, 1969; Howell, 1997; Thomas, Salazar, & Landers, 1991). The overall mean ($M = 19.36, SD = 2.52$) was more than 17.5. Therefore, based on the magnitude of the difference, the teachers in this study perceived authentic assessment use to have positive impact on students' self-concept.

Motivation

The result of the one-sample t-test on perceived impact on motivation was significant, $t(157) = 9.02, p < .05$. Estimate of Cohen's magnitude of the difference was large, $d = 0.72$. The total mean ($M = 19.27, SD = 2.46$) for motivation was greater than the neutral score of 17.5. In general, the physical education teachers in this study perceived authentic assessment has influenced students' motivation positively.

Skill Achievement

The result of the one-sample t-test on perceived impact on skill achievement was significant, $t(157) = 10.33, p < .05$. The magnitude of the difference was large, $d = 0.82$. The overall mean ($M = 19.44, SD =$

2.36) for skill achievement was above the neutral value of 17.5. This indicated that, the physical education teachers who use authentic assessment perceived it to have a positive influence on students' skill achievement.

Summary

Taken together the means, standard deviations, and the magnitude of the differences, the physical education teachers perceived authentic assessment use has made a positive impact on students' self-concept, motivation, and skill achievement. The rest of the analysis in this section of the results focuses on the investigation of the differences in the teachers' perceptions.

Differences in Perceptions about the Impact of Authentic Assessment use on Students' Self-Concept, Motivation, and Skill Achievement

In order to determine the appropriate statistical test to use to investigate differences in the physical education teachers' perceptions about the impact of authentic assessment use on students' self-concept, motivation, and skill achievement, the three subscales of the PEAAI (self-concept, motivation, and skill achievement) were entered into a bivariate correlation analysis. Correlations of the PEAAI subscales were moderate, $r = .42$ to $.63$ (Huck & Cormier, 1996). The correlations show that the three subscales measure some of the same qualities.

The intercorrelations of the PEAAI subscales suggested that the researcher use Multivariate Analysis of Variance (SPSS, 1995). Therefore, to investigate differences in the physical education teachers' perceptions about the impact of authentic assessment use, a 2 x 3 (gender x grade level) Multivariate Analysis of Variance (MANOVA) was calculated. The perceived impact scores were the dependent measures (see Table 5).

Table 5

Means and Standard Deviations of Perceived Impact of Authentic Assessment Use on Self-Concept, Motivation, and Skill Achievement by Gender by Grade Level

GENDER			Elm. (<u>n</u> = 71)	Mid/Jr. (<u>n</u> = 47)	High. (<u>n</u> = 40)
FEMALE	Self-Concept	<u>M</u>	19.73	20.00	19.47
		<u>SD</u>	(2.28)	(2.93)	(1.69)
	Motivation	<u>M</u>	19.78	19.04	19.05
		<u>SD</u>	(2.29)	(3.56)	(2.70)
	Achievement	<u>M</u>	19.41	19.79	19.41
		<u>SD</u>	(2.89)	(1.91)	(3.44)
MALE	Self-Concept	<u>M</u>	19.00	19.00	18.78
		<u>SD</u>	(2.76)	(2.33)	(2.82)
	Motivation	<u>M</u>	18.96	19.04	19.39
		<u>SD</u>	(2.02)	(2.22)	(2.03)
	Achievement	<u>M</u>	19.66	19.13	19.17
		<u>SD</u>	(1.84)	(2.00)	(1.80)

Note. No significant difference was found, $p > .05$.

Results of the 2 x 3 multivariate analyses showed no significant interaction effect for gender and grade level on perceived impact on self-concept, motivation, and skill achievement, Wilks Lambda, $\lambda = .96$, approximate $F(6, 300) = .89$, $p > .05$. In addition, no significant main effect was found for gender, $F(3, 150) = 1.63$, $p > .05$ or for grade level, $F(6, 300) = 0.36$, $p > .05$. Thus, female and male physical education teachers at all grade levels perceived a similar positive impact of authentic assessment use on students' self-concept, motivation, and skill achievement.

Interviews

Participants

In addition to the survey, 12 physical education teachers were interviewed. There were 6 females and 6 males in the interview sample. Four of the participants taught at the elementary school, 4 taught at the middle/junior high, and 4 taught at the high school level. The interview subjects were selected to ensure equal representation by gender and by grade level. All interviewees had participated in the survey.

Protocol

The interview participants were contacted over the telephone for permission to be interviewed. All the participants were asked to indicate the times of the day they teach and the times they wanted to be interviewed. The office address of each participant was requested. Each

participant was told the interview would not take more than fifteen minutes. To ensure that participants were available for the interview, each participant was re-contacted two days before the scheduled interview. All interviews were conducted during the spring 2000 semester, following observation. The observation, though not part of the research process, helped the researcher understand better the various activities that take place in contemporary physical education classrooms.

Questions Guiding the Interview

The purpose of this interview was to investigate further the context of physical education teachers' perceptions about the impact of authentic assessment use on students' and self-concept, motivation, and skill achievement. In order to investigate this purpose, these questions were asked in the following order:

1. How do you assess your students? Please give examples.
2. What do you call this type of assessment? For example authentic or performance or alternative assessment.
3. How long have you used this form of assessment?
4. Do you also use traditional assessment measures - paper and pencil tests?
5. Why do you use authentic or performance or alternative assessment measures?

6. Some advocates think authentic assessment has increased students' self-concept, motivation, and skill achievement. Do you think the use of authentic assessment has improved your students' self-concept, motivation, and skill achievement? Please give examples.

7. What assessment technique (s) did you use before using authentic assessment? Why did you bring authentic assessment measures into your assessment approach?

Data Collection

The researcher was at each interview site one hour before the interview. After the initial introduction, the physical education teacher and the researcher walked to the gymnasium for the class. Every teacher interviewed introduced the researcher to the students as an observer from the University of Northern Iowa. After the introduction, the researcher took a position in the gymnasium and watched the class. At the end of the class, the researcher walked with the physical education teacher to the office for the interview.

Each interviewee read, signed, and kept a copy of the informed consent form (see Appendix D) before the interview began. All interviews were tape recorded and transcribed. Although the observation was not recorded as part of the research process, it helped in contextualizing the interview and its results. In addition, it helped the researcher understand

better the various activities that take place in contemporary physical education classrooms.

Data Analysis

As Becher, Geer, Hughes, and Strauss (1961), Bogdan and Biklen (1982), and Miles and Huberman (1984) suggested, data analysis began with the very first data collected and was an ongoing process. Allowance was made from the beginning for possible emerging themes, which would explain the physical education teachers' perceptions about the impact authentic assessment use has made on their students. Furthermore, there was careful examination for clues to the teachers' reasons for using authentic assessment.

Emergent Themes: Teachers' Perceptions

The emergent themes represent ideas, beliefs, and concerns about authentic assessment use in the public school physical education programs. In some cases, the themes represent the rationales some teachers gave for using authentic assessment.

For clarity and identification, four emergent themes were labeled. The first theme--self-concept, motivation, and skill achievement--concerns physical education teachers' perceptions about the impact of authentic assessment use on students' self-concept, motivation, and skill achievement. The other three themes--effort and participation,

fitness and skill testing, and paper and pencil testing--help explain the types of assessment physical education teachers' use.

The Theme of Motivation, Self-Concept, and Skill Achievement

Every physical educator looks for ways to motivate students. According to Moursund (1976), "motivation is the wellspring of the learning process. Once a person is motivated, we say, the battle is all but won" (p. 299). Virtually every teacher interviewed mentioned motivation as a major component toward successful learning, and they believed authentic assessment motivates students because it gives them time to practice and perform better. The teachers indicated that students become motivated when they see positive things on their report cards. Accomplishing this in physical education requires a holistic assessment that never loses sight of the students. Such assessment should give students more than one attempt at the performance with feedback from the teacher. As said by one of the elementary teachers "students are pretty motivated in physical education but I think the assessment I use makes a difference." Although elementary students love physical activities and will try hard to succeed in physical education classes, teachers who use authentic assessment techniques in their programs reiterated that the techniques have enhanced students' motivation. In much the same way, one of the teachers used a project,

“Moving Across Iowa,” to illustrate the link between authentic assessment and motivation.

Moving across Iowa is a fitness project. In this project, the students’ walk, jog or walk/jog a seven-mile course from one point to the next until the whole course is completed. The course is divided into seven stations and must be completed in two weeks. Each station is named after a town in the state of Iowa. The students can complete the project on their own, with a sibling or with an adult. If they do with an adult, it doubles the points that the student gets.

Although challenging, students enjoyed the “Moving Across Iowa” project. Also, the students became physically fit and their motivation toward physical activity increased.

Furthermore, in physical education, there is a connection between motivation and performance. The connection between motivation and performance is so strong that for some teachers, motivating students to try harder has become an end in itself. Therefore, using different teaching methods and assessment techniques that help students succeed motivates them more. Thus, the teachers perceived that the use of authentic assessment motivates students to practice and improve performance.

As noted earlier, some of the emergent themes are interrelated, but in particular, motivation appears to be closely related to self-concept. By self-concept, teachers mean the feelings and attitudes students have about themselves (Woolfolk, 1990). Much of the teachers’ purpose for the use of authentic assessment was to motivate students to try harder

so improvement would occur. The teachers interviewed stated that what improves students' self-concept is that the students see and feel they have improved. However, that does not mean the lack of improvement implied a lack of motivation and lowering of self-concept. The teachers see improvement in students' self-concept when they accomplish challenging tasks, especially in the swimming pool.

We see much of students' self-concept in the pool most especially when a student masters parts of a swimming unit and can get to the deep end. Furthermore, when a beginning swimmer can float and perform basic swimming strokes so that she/he can compete with other students. Similarly, in the gymnasium, when a student shoots the basketball or bumps the volleyball at different levels so that she/he becomes skilled at it, we see improvement in self-concept.

Since public school physical education deals mostly with overt behaviors, students openly show their feelings and attitudes when they accomplish difficult tasks, and that becomes more obvious when teachers frequently use multiple techniques to assess students' behavior.

While instruction tends to be aimed at a group's mean level of ability, assessment appears to focus on individual achievement (Veal, 1988). The teachers in this interview emphasized that authentic assessment has improved students' skill achievement. When physical education teachers assess, they weigh the student's natural ability and past experiences with the amount of time spent in class. Physical education teachers who use authentic assessment no longer assess students' skills at the beginning and at the end of a course unit. Instead,

students are assessed during game situations and in relation to their level of ability. Such assessment technique gives students the chance to improve and better individual performances. Furthermore, it has enabled the teachers to detect improvement in both the highly skilled and those experiencing the activity for the first time. In addition, it has changed the grading practices of the teachers.

The Theme of Effort and Participation as Assessment Techniques in Physical Education

The physical education teachers in this study assess students on their effort. The teachers regard effort as internal and judge students' effort on participation, how they hustle in games, dressing, attendance in class, and working on skills instead of conversing with friends. The teachers judge these outward behaviors continuously and keep records of them. Also, effort plays a role in determining students' grades. For example, "a grade of 'A' means the student put in maximum effort." Thus, the less the effort a student put in a skill, the lower the grade. However, this does not mean skill is not considered when grading. Therefore, a low skilled student is assessed on his or her ability and effort and not just marked down.

Furthermore, a student's effort is judged by improvement. About one half of the teachers interviewed believed there is a link between effort and student learning. For example, one of the interviewees said:

I think the way we assess now gives students more chance to get good grades. At first it was more athletics because it was the skilled person that benefited. Now how we assess gives every student the chance to put in maximum effort to get a good grade.

To these teachers, a student's effort in most situations is equally important as the skill.

The Theme of Fitness and Skill Testing in Physical Education

Research in physical education has shown that assessment should not be used solely for the purpose of grading students, but to diagnose, correct, and improve performance (Hensley, 1997; Veal, 1988). The physical education teachers in this study used fitness test to diagnose and to report on students' fitness levels to the state. The fitness test commonly used by all the physical education teachers is the Presidential Fitness Test. With the Presidential Fitness Test, the students run a mile, shuttle run, v-sits, and do sit up or curl up. Some of the teachers do the Presidential Fitness test because it is mandatory and they would be held accountable if they do not report any data on their students. For those teachers, given the option, fitness testing would not form part of their physical education program. Others use the fitness test at the beginning of every semester. For those teachers who do fitness testing at the beginning of each semester, the fitness test gives them information about the fitness levels of their students.

As one teacher put it, "I do physical fitness testing at the beginning of every semester to see how good the students can run a mile, v-sits, pull ups, and

sit up. I then develop exercise programs for them individually.” When teachers in this study use fitness tests, it was for the purpose of accountability, diagnosing, and improvement of performance.

The teachers in this study rarely use skill tests. The teachers’ reasons for not using skill test vary. Some think that the conditions under which skill tests are administered differ from the real game situations. In addition, they think skill tests do not mimic the challenges of games. Others think assessing students on their skills wastes time. In addition, it reinforces the idea of preparing athletes instead of life long physically educated students. As commented by one of the teachers:

I personally do not think we should get into making students athletes by assessing students’ skills separately in our physical education classes. All I want is for my students to have fun and succeed in my class. Because of this I use a variety of assessment techniques.

Most physical education teachers interviewed felt that it was not only the skilled student that should get good grade. The teachers believed students have to understand the physical education concepts. Because of that, they broaden their assessment practices to include the affective domain in order not to favor only the skilled student. To these teachers, practicing skills on daily basis as part of the physical education program is more appropriate than assessing students’ skills in isolation.

The Theme of Paper and Pencil Testing

The teachers in this interview were divided on the use of paper and pencil tests. About a half of the physical education teachers interviewed use paper and pencil tests including short answer, multiple-choice tests, and true or false tests. According to these teachers, they use paper and pencil tests not to grade the students but to find out how much information the students have retained. For example, one of the teachers said, "I use paper and pencil tests to find out how much information the students have about playing the game of football or basketball."

The remaining half of the teachers did not use any paper and pencil test. According to these teachers' paper and pencil tests do not give accurate information about what the students can and cannot do. The teachers who do not use paper and pencil tests use different techniques including demonstration and event task to assess their students. As one of the teachers stated, "I assess my students more on their performance and participation because I think it is accurate and give parents more information about their child's performance."

Summary

Results of this study show that authentic assessment is used extensively in Iowa public school physical education. All the teachers in this study that use any authentic assessment generally use the same types of authentic assessment techniques. Most popular among them is

teacher observation. The teachers rely on observation to analyze and evaluate students' performances. In addition, the physical education teachers assess students on their effort and participation in class.

The physical education teachers in this assess students on their outward behaviors. The teachers use fitness tests to diagnose, remediate and improve performance. Furthermore, the teachers think using authentic assessment gives parents accurate information about their student's performance. Finally, the physical education teachers perceived authentic assessment use has positive impact on students' self-concept, motivation, and skill achievement.

CHAPTER V

DISCUSSION AND RECOMMENDATIONS

The discussion section of this study pulled together findings from both the mail questionnaire and the interview data. In this discussion section, possible reasons and explanations for the findings in this study are provided. Where applicable, quotations from the interviews are used to strengthen the discussion. In addition, tables from the survey analyses, and available literature are referenced to support the discussion. Finally, recommendations for further line of inquiry and practice are made. The purpose of this study was to describe and analyze the type and the extent of authentic assessment use in Iowa public school physical education. In addition, another purpose was to investigate physical education teachers' perceptions of the impact of authentic assessment use on students' self-concept, motivation, and skill achievement.

Discussion

The discussion is organized into the following headings and subheadings:

1. Extent of authentic assessment use including the use of authentic assessment techniques among “users, and none “users.”
2. Perceived impact about authentic assessment use on students' self-concept, motivation, and skill achievement.

Extent of Authentic Assessment Use

In many ways, the findings of this study are not surprising. Other investigators including Hensley (1997) and Melograno (1994) have shown that alternative assessment is not new to the physical education profession. What is interesting in this study is the extensive use of authentic assessment in public school physical education. The extent of use indicated that 75.2% of physical education teachers in this study use authentic assessment techniques. One possible reason for the teachers' extensive use of authentic assessment may be that the teachers were exposed to authentic assessment techniques in their teacher education programs. The subjects' demographic information revealed that the teachers in this study ranged in physical education teaching experience from 1 year to 35 years; with 17.0 being the median. With this teaching experience, it is possible that majority of the teachers have used authentic assessment practices over the years. Therefore, it is not surprising when three-quarters of the teachers indicated they use authentic assessment technique. Again, research has shown that alternative assessment is not new to the physical education profession (Weinberg, 1996).

A possible explanation for the extensive use of authentic assessment is the very practical nature of public school physical education. Existing literature, for example, Placek (1983) and

Tousignant and Siedentop (1983), indicates that public school physical education teachers are primarily concerned with students' outward behavior. The findings in this study support this early research as one of the teachers interviewed said, "I assess my students on their natural abilities, how they perform in class, and their contribution in game situations." For physical education teachers to be able to assess students on these outward behaviors requires the use of authentic assessment techniques including teacher observation, self-observation, demonstration, peer observation, event tasks and checklist.

Another explanation for the teachers extensive use of authentic assessment may be the emphasis placed on the use of authentic assessment techniques in the 1995 National Physical Education Standard (NASPE) publication, *Moving into the Future* (NASPE, 1995). The guidelines in this book recommended that physical educators use different authentic assessment techniques to capture students' learning. As the physical education teachers in this study generally rely on the NASPE guidelines to meet the demands of their programs, it is possible that the suggested assessment techniques in the NASPE guidelines may have influenced their assessment practices.

An additional plausible explanation for the extensive use of authentic assessment may be the increase of accountability in the public schools. Today, public school physical education teachers are required

to provide data on students' fitness levels and information on students' progress to the state and parents, respectively. The following is a quote from one of the teachers: "I have to gather these fitness data so that my boss can forward them to the state department of education. If I don't get them ready on time, I will be in trouble with my boss." In order to meet the demands of accountability in the public school and at the same time provide direct evidence of students' skill achievement to parents, physical education teachers rely on authentic assessment techniques (Lund, 1997).

Use of Authentic Assessment Techniques Among "Users"

All the physical education teachers who use any authentic assessment technique use teacher observation (see Tables 1 and 2). One factor which might lead to the consistent use of teacher observation is that physical education classes in the public school systems deal with overt behaviors, and focus on student learning and active participation (Lund, 1997; Siedentop, 1991). The overt behaviors exhibited by students in physical education classes make teacher observation a logical choice for assessing students' performance (Hensley, 1997).

In addition, many physical performances are fast paced and no single assessment technique can capture all the component parts. Therefore, the teachers rely on observation and professional knowledge to assess students' performances.

The findings on the extent of use in this study showed that a higher proportion of elementary teachers reported using checklists than do high school teachers (see Table 2). A plausible explanation may be that elementary physical education teachers are more concerned with the development and documentation of students' fundamental skills and progress than the high school teachers. The higher proportion of checklist use at the elementary level supports Gronlund and Linn (1990) who indicated that checklist use is one of the useful ways to record students' progress especially in the elementary level.

However, the frequency of use of checklists indicates that high school teachers use checklists more often than elementary teachers do (see Table 3). A possible explanation for the high school teachers' frequent use of checklists may be the reduction in the number of physical education teachers at this grade level. All the four high school teachers interviewed and those who indicated in the mail survey that they did not use any authentic assessment techniques complained about the reduction in their teaching staff. In addition, these teachers lamented about the increase in the number of students in their classes. Taken together the reduction in the number of teaching staff in the high school physical education programs and the increase in class sizes, the high school teachers' frequent use of checklists is justified. Research

indicates that checklists are one of the convenient means of recording judgments on students' performances (Linn & Gronlund, 1995).

The essay and the portfolio were the least used authentic assessment techniques. The proportion of use (see Tables 1 and 2) indicated that nearly half of the physical education teachers in this study use these assessment techniques. According to Lund (1997), essays are used to describe and explain facts in context. Portfolios, on the other hand, are collections of students' work over time (Melograno, 1994; 2000). Both the essay and the portfolio are authentic assessment techniques that require time to implement. In addition, the number of contact days in some public school physical education has been reduced. The given reasons make it difficult for the physical education teachers to use the essay and the portfolio assessment techniques especially where they see students not more than twice in a week.

Another possible explanation for the limited use of the portfolio may be that portfolio is relatively new in physical education (Killoran, 1992). Because the portfolio technique is relatively new in physical education, it is possible that majority of the teachers who use authentic assessment did not have the opportunity to use the portfolio technique in their teachers' preparation programs. Hence, they felt uncomfortable using it in their physical education classrooms.

The findings in this study showed that female and male physical education teachers at all grade levels did not differ significantly on their use of the popular authentic assessment techniques (see Tables 3 and 4). Perhaps the regular use of the popular authentic assessment techniques has become part of the physical education practice. Also, there is progression and continuity in the physical education curriculum. The middle/junior high school physical education teachers take over from the elementary teachers and build on the fundamentals they laid. The same occurs in the high school physical education program. The continuity in the physical education curriculum might have resulted in the consensus in the frequency of use of the popular authentic assessment techniques at all grade levels.

Non-“Users” of Authentic Assessment Techniques

While three quarters (75.2%) of the public school physical education teachers in this study use authentic assessment, the remaining one-quarter (24.8%) do not. About 29% of the teachers who indicated that they did not use authentic assessment said they did not because “authentic assessment consumes time.” Available literature indicates that authentic assessment requires time to plan, implement, and evaluate (Kirk, 1997; Lund, 1997). The time-consuming nature of authentic assessment might have become a barrier to these teachers’ use of authentic assessment techniques. Also, in most public school

physical education programs, the number of teachers has been reduced.

All the high school teachers who indicated that they did not use authentic assessment stated that the number of teachers in their programs has been reduced. For example, one of the high school teachers said "I cannot do much these days because we are only three teachers here and I have so many students in my class. I try to accomplish my goals for the day but it is just like get them in and get them out." The reduction in the number of physical education teachers especially in the high school may be one of the explanations why some teachers do not use authentic assessment.

Furthermore, some public school physical education teachers often operate under difficult circumstances and many problems exist which prevent them from utilizing all the techniques they know (Veal, 1988). For example, all the teachers who indicated that they did not use authentic assessment reported they teach every hour of the school day and supervise students during lunch. The problems physical education teachers face in their schools become magnified when students come to class only one or two times in a week, and in big class sizes. Taken together, the above reasons suggested and the complaints, the teachers may be justified in not using authentic assessment techniques in their programs.

Perceived Impact About Authentic Assessment Use on Students' Self-
Concept, Motivation, and Skill Achievement.

Another factor to consider is the nature of the physical education teachers' perceptions about the impact authentic assessment use has made on students. All the physical education teachers in this study perceived authentic assessment to have positive influences on students' self-concept, motivation, and skill achievement.

A possible explanation for the perceived positive impact may be that physical education students approach the physical education program with different attitudes, motivations and skill levels. As the physical education teachers use different teaching methods and assessment techniques that enable students' demonstrate their competence, they see positive changes and improvements in students' self-concept, motivation, and skill achievement. The teachers interviewed stated that they see an increase in these attributes in the gymnasium and in the swimming pool more especially when the students accomplish challenging tasks.

An additional possible explanation may be the consistent use of authentic assessment techniques in the public school physical education programs. The physical education teachers in this study generally use the same assessment techniques regardless of gender and grade level (see Tables 1 and 2). As the students go through each phase of the

program, they become accustomed to the teachers' assessment practices. The learning effect experienced by the students through the continuous use of the assessment techniques enhances their self-confidence and motivation. Thus, the physical education teachers' see these attributes manifest in the students when the same assessment techniques are used.

Recommendations for Research

Authentic assessment is used extensively in Iowa public school physical education. In addition, the physical education teachers who use authentic assessment perceived it has positive influence on students' self-concept, motivation, and skill achievement. However, before closing the issue on this topic, more research is needed to broaden our understanding of authentic assessment use and how it influences students' behavior in and out of school. The following recommendations are for further investigation:

1. Given that this study used survey technique to collect data, a pure qualitative research using observation and in-depth interviews may help understand the types of authentic assessment technique used in contemporary public school physical education.
2. Given that the Physical Education Authentic Assessment Inventory (PEAAI) used in this study was not fully validated, a study is needed to validate the instrument for its continue use.

3. Given that this study focused on teachers' perceptions, further research investigating students' attitude about the use of authentic assessment is recommended.

4. Given that about one-quarter of the physical education teachers in this study did not use any authentic assessment technique, research is needed to investigate what assessment these non- "users" of authentic assessment use.

5. Research is needed to investigate if using authentic assessment causes an increase in students' self-concept, motivation, and skill achievement.

6. Finally, research is needed to understand the unique character of physical education goals and subject matter and its relationship with authentic assessment use.

Recommendations for Practice

Based on the findings in this study, it is recommended that:

1. Public school physical education programs get adequate resources and national support.

2. Physical education teacher programs should prepare students to use different authentic assessment techniques including teacher observation.

REFERENCES

- Alcove, L. D., & McCarty, B. J. (1992). Plain talk: Recognizing positivism and constructivism in practice. Action in Teacher Education, 14(2), 16-22.
- Archbald, J., & Newman, F. M. (1988). Beyond standardized testing: Assessing authentic academic achievement in the secondary school: Reston, VA: National Association of Secondary School Principals.
- Becker, H. S., Geer, B., Hughes, E. C., & Strauss, A. L. (1961). Boys in white, student culture in medical school. Chicago: University of Chicago Press.
- Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain. New York: Longman.
- Bodner, G. M. (1986). Constructivism: A theory of knowledge. Journal of Chemical Education, 63(10), 873-877.
- Bogdan, B., & Biklen, T. (1982). Qualitative research for education: An introduction to theory and methods. Boston: Allyn & Bacon.
- Brooks, J. G., & Brooks, M. G. (1993). In search of understanding: The case for constructivist classrooms. Alexandria, VA: Association for Supervision and Curriculum Development.
- Caine, R. N., & Caine, G. (1997). Education on the edge of possibility. Alexandria, VA. Association for the Supervision and Curriculum Development.
- Chittenden, E. (1991). Authentic assessment evaluation, and documentation of student performance. In V. Perrone (Ed.), Expanding student assessment (pp. 22-32). Alexandria, VA: Association for Supervision and Curriculum.
- Cizek, G. J. (1991). Innovation or enervation? Performance assessment in perspective. Phi Delta Kappan, 72(9), 695-699.

- Cleland, F. E., & Stevenson, R. (1997). Authentic assessment in action. Teaching Secondary Physical Education, 3(3), 4-7.
- Cobb, P., & Steffe, L. P. (1983). The constructivist researcher as teacher and molder builder. Journal for Research in Mathematics Education, 14(2), 83-94.
- Cohen, J. (1969). Statistical power analysis for the behavioral sciences (2nd ed.). New York: Academic Press.
- Cole, N. (1992). Changing assessment practice in mathematics education. Reclaiming assessment for teaching and learning. Paper presented at the Conference on Partnership for Systemic Change in Mathematics, Science, and Technology Education, Washington, DC.
- Converse, J. M., & Presser, S. (1986). Survey questions handcrafting the standardized questionnaire. Beverly Hills, CA: Sage Publications, Inc.
- Dana, T. M., & Tippins, D. J. (1993). Considering alternative assessments for middle learners. Middle School Journal, 25(2), 3-5.
- Darling-Hammond, L. (1990). Setting standards for students: The case for authentic assessment. NASSP Bulletin, 19-26.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. Harvard Educational Review, 64(1), 5-26.
- DeVries, R. (1987). Constructivist education: Implications of Piaget's theory. In N. Hazareesingh & J. Livingston (Eds), Early education: Building bridges to the future (pp. 6-17), Duluth, MN: University of Minnesota.
- Diez, M. E., & Moon, J. C. (1992). What do we want students to know?...and other important questions. Educational Leadership, 49(8), 38-39.
- Dunham, P., Jr. (1994). Evaluation for physical education. Marton, Co: Englewood Publishing Company.

- Ellis, A. K., & Fouts, J. T. (1993). Research on educational innovations. Princeton Junction, NJ: Eye on Education, Inc.
- Feuer, M. J., & Fulton, K. (1993). The many faces of performance assessment. Phi Delta Kappan, 74(6), 478.
- Glesne, C., & Peshkin, A. (1992). Becoming qualitative researchers: An introduction. White Plains, NY: Longman.
- Gronlund, M. A., & Linn, J. E. (1990). Measurement and evaluation in teaching (6th ed.). New York: Macmillan Publishing Company.
- Haney, W., & Madaus, G. (1989). Searching for alternatives to standardized tests: Whys, whats, and whithers. Phi Delta Kappan, 70(9), 683-687.
- Hauge, L. S. (1997). Authentic assessment: Utilizing portfolios in physical education. California Association for Health, Physical Education, Recreation, and Dance, 59(4), 23-24.
- Head, D. (1996). The plain truth about authentic assessment. California Association for Health, Physical Education, Recreation and Dance, 59(2), 5-6.
- Hensley, L. (1997). Alternative assessment for physical education, Journal of Physical Education, Recreation, and Dance, 68(7), 19-24.
- Hensley, L., Aten, R., Baumgartner, T. A., East, W. B., Lambert, L. T., & Stillwell, J. L. (1989). A survey of grading practices in public school physical education. Journal of Research and Development in Education, 22(4), 37-42.
- Herman, J. L., Aschbacher, P. R., & Winters, L. W. (1992). A practical guide to alternative assessment. Alexandria, VA: Association for Supervision and Curriculum Development.
- Herman, J. L., & Winters, L. (1994). Synthesis of research. Portfolio research: A slim collection. Educational Leadership, 52(2), 48-54.
- Holmes, G. A., & Leitzel, T. C. (1993). Evaluating learning through a constructivist paradigm. Performance and Instruction, 32(8), 28-30.

- Howell, D. C. (1997). Statistical methods for psychology. Belmont, CA: Wadsworth Publishing Company.
- Huck, S. W., & Cormier, W. H. (1996). Reading statistics and research. New York: Haper Collins College Publishers Inc.
- Imwold, C. H., Rider, R. A., & Johnson, D. J. (1982). The use of evaluation in public school physical education programs. Journal of Teaching Physical Education, 2(1), 13-18.
- Jones, D. H. (1993). Using authentic assessment in elementary social studies. Social Science Record, 30(1), 17-24.
- Katims, N., Nash, P., & Tocci, C. M. (1993). Linking instruction and assessment in a middle school mathematics classroom. Middle School Journal, 25(2), 28-35.
- Killoran, J. (1992). In defense of the multiple-choice question. Social Education, 56(2), 107-109.
- Kirk, M. F. (1997). Using portfolios to enhance student learning and assessment. Journal of Physical Education, Recreation, and Dance, 68(7), 29-33.
- Krathwohl, D. R. (1993). Methods of educational and social science research. An integrated approach. White Plains, NY: Longman Publishing Group.
- Kritt, D. (1993). Authentic, reflection, and self-evaluation in alternative assessment. Middle School Journal, 25(2), 43-45.
- Lincoln, Y., & Guba, E. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage.
- Linn, R. L., & Gronlund, N. E. (1995). Measurement and assessment in teaching (7th ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Lund, J. (1997). Authentic assessment: Its development and applications. Journal of Physical Education, Recreation, and Dance, 68(7), 25-28.
- Madaus, G. F., & Kellaghan, T. (1993). The British experience with authentic testing. Phi Delta Kappan, 74(6), 458-469.

- Manus, A. L. (1996). Procedural versus constructivist education: A lesson from history. The Educational Forum, 60, 312-316.
- Marzano, R. J., Pickering, D., & McTinghe, J. (1993). Assessing student outcomes: Performance assessment using the dimensions of learning model. Alexandria, VA: Association of Curriculum Development.
- Melograno, V. J. (1994). Portfolio assessment: Documenting authentic student learning. Journal of Physical Education, Recreation, and Dance, 65(8), 50-55, 58-61.
- Melograno, V. J. (2000). Designing a portfolio system for K-12 physical education: A step-by-step process. Measurement in Physical Education and Exercise Science, 4(2), 97-115.
- Meyer, C. A. (1992). What's the difference between authentic assessment and performance assessment? Educational Leadership, 49(8), 39-41.
- Meyer, C., & Schuman, S. (1990). White paper on aggregating portfolio data. Oswego, OR: Northwest Evaluation Association.
- Miles, M. B., & Huberman, A. M. (1984). Qualitative data analysis: A sourcebook of new methods. Beverly Hills, CA: Sage Publications.
- Mitchell, R. (1992). Testing for learning: How new approaches to evaluation can improve American schools. New York: The Free Press.
- Moursund, J. P. (1976). Learning and the learner. Belmont, CA: Wadsworth Publishing Company, Inc.
- National Association for Sport and Physical Education. (1995). Moving into the future. St. Louis, MO: Mosby.
- National Commission on Excellence in Education. (1983). A nation at risk: The imperative of educational reform. Washington, DC: National Commission on Excellence in Education.
- Neil, D. M., & Medina, N. J. (1989). Standardized testing: Harmful to educational health. Phi Delta Kappan 70(9), 688-696.

- Noel, J. R. (1993). Constructivist theory and practice in teacher education. Paper presented at the 1993 Annual Meeting of the American Educational Research Association, Atlanta, GA.
- Novak, J. D., & Gowin, D. B. (1984). Learning how to learn. Cambridge, MA: Cambridge University Press.
- O'Neil, J. (1992). Putting performance assessment to the test. Educational Leadership, 49(8), 14-19.
- Pankratius, W. J., & Young, M. W. (1995). Perspectives on education: A constructivist approach to an introductory course. Education, 115(3), 363-370.
- Parker, M. (1998). Authentic assessment in physical education: A few ideas to use tomorrow. Paper presented at CDAAHPERD, February 13-15, Wichita, KS.
- Paulson, F. L, Paulson, P. R., & Meyer, C. A. (1991). What makes a portfolio a portfolio? Educational Leadership, 48(5), 61-63.
- Placek, J. M. (1983). Conceptions of success in teaching: Busy, happy, good? In T. J. Templin & J. K. Olsen (Eds.), Research on teaching in physical education (pp. 45-56). Champaign, IL: Big Ten Symposium.
- Prawat, R. S. (1992). Teachers' beliefs about teaching and learning: A constructivist perspective. American Journal of Education, 100(3), 354-395.
- Powell, J. C. (1993). What does it mean to have authentic assessment? Middle School Journal, 25(2), 36-45.
- Pullin, D. C. (1994). Learning to work: The impact of curriculum and assessment standards on educational opportunity. Harvard Educational Review, 64(1), 31-52.
- Rea, L. M., & Parker, R. A. (1992). Designing and conducting survey research. San Francisco: Jossey-Bass Publishers.
- Rossi, P. H., Wright, J. D., & Anderson, A. B. (1989). Handbook of survey research. Orlando, FL: Academic Press, Inc.

- Ryan, J. M., & Miyasaka, J. R. (1995). Current practices in testing and assessment. What is driving the changes? Bulletin.
- Siedentop, D. (1991). Developing teaching skills in physical education (3rd ed.). Mountain View, CA: Mayfield.
- Statistical Package for Social Sciences for Windows, Release 7.0. (1995), Computer software. Chicago: Statistical Package for Social Sciences Inc.
- Thomas, J. R., Salazar, W., & Landers, D. M. (1991). What is missing in $p < .05$? Effect size. Research Quarterly for Exercise and Sport, 62(3), 344-348.
- Tinsley, C. (1993). Putting first things first: Changing our assessment practices. Middle School Journal, 25(2), 51-54.
- Tousignant, M., & Siedentop, D. (1983). A qualitative analysis of task structures in required secondary physical education classes. Journal of Teaching in Physical Education, 3(1), 47-57.
- Veal, M. L. (1988). Pupil assessment perceptions and practices of secondary teachers. Journal of Teaching Physical Education, 7, 327-342.
- Von Glasersfeld, E. (Ed.). (1990). Radical constructivism in mathematics education. Dordrecht: Kluwer.
- Vygostky, L. (1986). Thought and language. Cambridge, MA: MIT Press.
- Weinberg, H. (1996). Authentic assessment in middle school physical education. Teaching Secondary Physical Education, 2(2), 20-21.
- Wheatley, G. H. (1991). Constructivist perspectives on science and mathematics learning. Science Education, 75(1), 9-12.
- Wiggins, G. (1989a). "Teaching to the (Authentic) Test." Educational Leadership, 46(7), 41-47.
- Wiggins, G. (1989b). A true test: Toward more authentic and equitable assessment. Phi Delta Kappan, 70(9), 703-713.

- Wiggins, G. (1993). Assessment to improve performance, not just monitor it: Assessment reform in the social sciences. Social Science Record, 30(2), 5-12.
- Wimmer, R. D., & Dominick, J. R. (1997). Mass media research An introduction. Belmont, CA: Wadsworth Publishing Company.
- Winn, W. D. (1991). The assumptions of constructivist and instructional design. Educational Technology, 31(9), 38-40.
- Winograd, P., Jones, D., & Perkins, F. (1994). The politics of portfolios, performance events and other authentic assessments. Paper presented at the Association 39th Annual Meeting of International Reading, Toronto, Ontario, Canada.
- Wood, T. M. (1996). Evaluation and testing: The road less traveled. In S. J. Silverman & C. D. Enins (Eds.), Student learning in physical education: Applying research to enhance instruction. Champaign, IL: Human Kinetics.
- Woolfolk, A. E. (1990). Educational psychology. Englewood Cliffs, NJ: Prentice Hall.
- Worthen, B. R. (1993). Critical issues that will determine the future of alternative assessment. Phi Delta Kappan, 74(6), 444-454.
- Zhu, W. (1997). Alternative assessment: What, why, how. Journal of Physical Education, Recreation, and Dance, 68 (7), 17-18.

APPENDIX A

TEACHER INVITATION LETTER

Dear Physical Educator:

I am contacting you to participate in a statewide survey. The purpose of this survey is to find out your perspective on the use of authentic assessment techniques in physical education. I am defining authentic assessment as any assessment task that uses multiple scoring systems to measure students' habits and repertoires related to life outside the classroom. Your responses will help colleges and universities in the state of Iowa to reconsider the preparation of physical educators in the state. Your name appeared in a scientifically selected random sample. Your answers are very important to the accuracy of this study even if you do not use any authentic assessment technique.

It will take only 10 minutes of your time to answer the simple questions on the enclosed questionnaire. There are no correct or incorrect responses, only your much-needed opinions and frequency of use are needed. This questionnaire has an identification number that will be used for follow-up purposes only. All responses are confidential and will be used only in combination with those of other physical educators once the survey process has been concluded.

Please complete the enclosed questionnaire. Seal it in the postage-paid, preaddressed envelope and drop in the mail by February 25, 2000. If you are interested in receiving a report on the findings of this research, write on the back of the questionnaire. I will be glad to send you a complimentary report when it is ready.

For answers to questions about the research project contact the office of the Human Subjects Coordinator, University of Northern Iowa, (319-273-2748) or the advisors overseeing this project, Dr. Susann Doody (319-273-2011) or Dr. Robert Boody (319-273-6198). To contact me the researcher, call Joseph Mintah (319-222-6035).

Sincerely,

Joseph K. Mintah
University of Northern Iowa

APPENDIX B

TEACHER INFORMATION

GENDER	Female	Male	
EDUCATIONAL LEVEL	Bachelor	Bachelor+ Masters	
	Masters+	Specialist Doctorate	
GRADE LEVELS TEACHING	Elementary	Middle/Jr. High	High School
TOTAL NUMBER OF YEARS TAUGHT	_____		
NUMBER OF YEARS TEACHING P. E.	_____		
YEARS IN PRESENT SCHOOL	_____		

For the purpose of this study authentic assessment is defined as assessment tasks that use multiple scoring systems to measure students' habits and repertoires on significant tasks related to life outside the classroom. This definition includes but is not limited to the following assessment techniques: teacher observation of student performance, portfolios, student projects, written essays, student exhibitions, and peer observation.

Do you presently use any authentic assessment techniques in your physical education classes?

_____ Yes, approximately how long have you used authentic assessment?

_____ Less than 1 year _____ 1-2 years _____ 3-5 years
 _____ 6-10 years _____ more than 10 years

If you use authentic assessment, complete sections B and C of the questionnaire.

_____ No, state reasons why authentic assessment is not used.

If you responded no above, you have finished answering the questionnaire. Thank you.

APPENDIX C

SECTION A

AUTHENTIC ASSESSMENT TECHNIQUES

PLEASE CIRCLE A RESPONSE FROM THE SCALE BELOW TO SHOW HOW FREQUENT YOU USE EACH OF THE ASSESSMENT TECHNIQUES AND CIRCLE THE APPROPRIATE MARK TO THE RIGHT OF EACH ITEM.

A = ALWAYS

O = OFTEN

S = SOMETIMES

R = RARELY

N = NEVER

	<u>A</u>	<u>O</u>	<u>S</u>	<u>R</u>	<u>N</u>
1. Teacher Observation	A	O	S	R	N
2. Peer Observation	A	O	S	R	N
3. Self-observation	A	O	S	R	N
4. Checklists	A	O	S	R	N
5. Event Task	A	O	S	R	N
6. Student Log	A	O	S	R	N
7. Demonstration	A	O	S	R	N
8. Written Essays	A	O	S	R	N
9. Portfolios	A	O	S	R	N
10. Parental Report	A	O	S	R	N
11. Students Projects	A	O	S	R	N
12. Group Projects	A	O	S	R	N
13. Anecdotal Records	A	O	S	R	N

A = ALWAYS		O = OFTEN		S = SOMETIMES		
R = RARELY		N = NEVER				
		<u>A</u>	<u>O</u>	<u>S</u>	<u>R</u>	<u>N</u>
14.	Students Projects	A	O	S	R	N
15.	Group Projects	A	O	S	R	N
16.	Anecdotal Records	A	O	S	R	N
17.	Video	A	O	S	R	N
18.	Oral Discourse	A	O	S	R	N
19.	Others	A	O	S	R	N
	(Indicate frequency of use)	A	O	S	R	N
		A	O	S	R	N

SECTION B

PHYSICAL EDUCATION TEACHERS' PERCEPTIONS

PROPONENTS OF AUTHENTIC ASSESSMENT HAVE MADE VARIOUS CLAIMS ABOUT ITS EFFECTIVENESS WHILE CRITICS HAVE CHALLENGED ITS USEFULNES. I AM INTERESTED IN YOUR OPINION. BELOW ARE STATEMENTS ABOUT AUTHENTIC ASSESSMENT. PLEASE INDICATE THE EXTENT TO WHICH YOU AGREE OR DISAGREE WITH EACH STATEMENT BY CIRCLING THE APPROPRIATE MARK TO THE RIGHT OF EACH STATEMENT.

SA = STRONGLY AGREE

A = AGREE

D = DISAGREE

SD = STRONGLY DISAGREE

	<u>SA</u>	<u>A</u>	<u>D</u>	<u>SD</u>
1. Students' achievement has improved as a result of using authentic assessment.	SA	A	D	SD
2. Students appear to enjoy class more since using authentic assessment.	SA	A	D	SD
3. As a result of authentic assessment students appear less confident in class.	SA	A	D	SD
4. As a result of authentic assessment students are more willing to try new things	SA	A	D	SD
5. Students' skill level in game situations has dropped as a result of authentic assessment	SA	A	D	SD
6. Students appear to find authentic assessment tasks enjoyable.	SA	A	D	SD

SA = STRONGLY AGREE		A = AGREE			
D = DISAGREE		SD = STRONGLY DISAGREE			
		SA	A	D	SD
7.	Students tend to push themselves toward higher goals as a result of authentic assessment.	SA	A	D	SD
8.	As a result of authentic assessment students appear more willing to take challenges.	SA	A	D	SD
9.	As a result of authentic assessment students' physical fitness has decreased.	SA	A	D	SD
10.	As a result of authentic assessment students utilize skills better in games.	SA	A	D	SD
11.	Students' motivation toward learning has decreased as a result of authentic assessment.	SA	A	D	SD
12.	Authentic assessment has had a positive effect on students' self-concept.	SA	A	D	SD
13.	Students feel negative about themselves as a result of authentic assessment.	SA	A	D	SD
14.	Students' performance in problem solving situations has improved.	SA	A	D	SD
15.	Students' participation in class has decreased as a result of authentic assessment.	SA	A	D	SD

Please write your comments pertaining to the effect of using authentic assessment on your students.

End of the questionnaire. Thank you very much.

APPENDIX D

INTERVIEW INFORMED CONSENT FORM

Dear Participant,

This study proposes to help our understanding of physical education teachers' perceptions about the impact authentic assessment use has made on students' self-concept, motivation, and skill achievement. Authentic assessment is any assessment task that uses multiple scoring systems to measure students' habits and repertoires and is related to life outside the classroom. If you agree to participate, I will request to interview you once.

There is no compensation for your participation and you have the right to refuse to answer questions and/or withdraw from the study at anytime. Transcripts (written copies) of my field notes from the interview and recorded conversation will be kept for my analysis. No information that could identify you will appear in the record. You will have the right to review the transcripts and listen to the recorded conversation at the end of the interview.

For answers to questions about the research project or the rights of research participants, you may contact the University of Northern Iowa, (319-273-2748) or Dr. Susann Doody, (319/273-2011) or Dr. Robert Boody (319/273-6198). To contact me, the researcher, call Joseph Mintah (319/222-6035).

I am fully aware of the nature and extent of my participation in this project as stated above. I agree to participate in this project. I acknowledge that I have received a copy of this consent statement.

(Signature of participant)

Date

(Printed name of participant)

(Signature of Investigator)

Date