

Think-Aloud Research Study

Analysis of Arizona Educator Proficiency Assessment (AEPA) verbal data

Vibeke Hjelseth



Institute for Educational Research

Faculty of Education

UNIVERSITY OF OSLO

May 2009

Abstract

We investigated the cognitive process of 45 beginning students, advanced students and experienced teachers who responded to selected items from Arizona Educator Proficiency Assessment (AEPA) practise tests while they were thinking aloud, AEPA is a licensure test required of all applicants for a teaching certificate in Arizona. Our attempt was to understand the types of mental processes used in taking these kinds of tests. We wanted to discover which mental process was the most frequent, if there were any mental processes related to performance on the test and if there were any differences between the three groups attending our study. Deciding about option was the most frequent thought process among the participants. Common sense was the only significant positive predictor related to test score. Error was a significant negative predictor in predicting test scores, which demonstrate that higher performers tended to rely less on test irrelevant strategies. Further we found differences between the three groups' mental thought processes. Advanced students tended to rely more on processes based on learning theories and research, experienced teachers drew more on processes based on own experience, and beginning students had a higher proportion of mental thought processes based on error.

Acknowledgements

Finally this process has come to an end. It has been an exciting but sometimes difficult process.

First of all I want to thank my professor in Arizona, Jerome D'Agostiono who let be a part of this project. I am grateful for this insight to the American educational system. Furthermore I would like to express thanks to my supervisor Helge Strømsø for his critical eye, encouragement and correction. He has been important in the completion of this thesis.

I would also like to show gratitude to Heidi for being such a good travel partner and for helping me with the language. I will also thank my neighbours Gro and Finken for helpful computer tips.

Oslo, May 2009

Vibeke Hjelseth

Contents

| | |
|--|-----------|
| ABSTRACT | 2 |
| ACKNOWLEDGEMENTS..... | 4 |
| 1. INTENTION | 9 |
| 1.1 STRUCTURE..... | 10 |
| 2. THEORY AND RESEARCH | 11 |
| 2.1 HISTORICAL BACKGROUND | 11 |
| 2.2 TEACHERS' MENTAL THOUGHT PROCESSES | 11 |
| 2.3 COMPLEXITY OF TEACHING | 12 |
| 2.4 THE DECISION MAKING PROCESS | 12 |
| 2.5 WHICH MENTAL THOUGHT PROCESSES ARE MOST FREQUENTLY USED WHILE TEACHERS ARE TO MAKE DECISIONS ABOUT HOW TO TEACH? | 13 |
| 2.6 ARE THERE ANY COHERENCE BETWEEN TEACHERS' MENTAL THOUGHT PROCESSES AND ACHIEVEMENT? | 15 |
| 2.7 ARE THERE ANY DIFFERENCES BETWEEN NOVICE AND EXPERIENCED TEACHERS' MENTAL THOUGHT PROCESS?..... | 16 |
| 2.8 EXPECTATION..... | 18 |
| 3. METHODOLOGICAL APPROACHES..... | 21 |
| 3.1 THE PURPOSE OF THE STUDY | 21 |
| 3.2 ANALYSES OF VERBAL DATA..... | 22 |
| 3.3 ADVANTAGES OF ANALYSIS OF VERBAL DATA | 24 |
| 3.4 LIMITATIONS OF ANALYSIS OF VERBAL DATA..... | 25 |
| 4. MATERIAL | 28 |
| 4.1 LICENSURE TESTS | 28 |
| 4.2 ARIZONA EDUCATOR PROFICIENCY ASSESSMENT (AEPA)..... | 29 |

| | | |
|------------|---|-----------|
| 4.3 | CONCERNS REGARDING LICENSURE TESTS | 29 |
| 4.4 | VALIDITY EVIDENCE OF LICENSURE TESTS | 30 |
| 5. | METHOD..... | 33 |
| 5.1 | PARTICIPANTS | 33 |
| 5.2 | PROCEDURE | 35 |
| 5.3 | PREPARING THE VERBAL REPORTS FOR ANALYSIS | 35 |
| 6. | RESULTS..... | 41 |
| 7. | DISCUSSION..... | 50 |
| 7.1 | LIMITATION | 53 |
| 8. | CONCLUSION..... | 55 |
| 9. | REFERENCES | 56 |
| 10. | APPENDIX | 61 |

List of tables

| | |
|--|----|
| Table 1: Teaching experience | 34 |
| Table 2: System for coding | 37 |
| Table 3: Descriptive statistic on cognitive processes | 42 |
| Table 4: Component Matrix..... | 43 |
| Table 5: Summary results of teaching level as a component..... | 46 |
| Table 6: Summary of regression analysis for variables predicting AEPA test score . | 48 |

1. Intention

An essential purpose for this research paper is to focus on the way teachers themselves understand teaching and their own roles in it. The thinking, planning and decision making of teachers represent a large part of the psychological context of teaching. Why do teachers teach the way they do? What are the actions an expression of? Some acts of practice are intentionally considered, others are routines, while others are intuitive.

This paper is an attempt to understand teachers' mental thought processes and if their theoretical understanding reflects their action in the classroom. In order to understand teachers' thoughts, judgments, decisions and behavior, one must be aware of how thoughts get carried into actions. One of the most important mental thought processes teachers are involved in is the decision making process. Teachers have to make decisions all the time. Research has discovered that teachers encounter decision situations at two minutes intervals while teaching (Peterson and Clark, 1986). Consequently, this paper also focuses a great deal on the decision making process.

In order to investigate teachers' mental thought processes, a group of four people from The University of Arizona asked 45 teachers and students to think aloud while answering to items taken from Arizona Educator Proficiency Assessment (AEPA). The test we used is supposed to measure basic skills, general knowledge, content knowledge, and knowledge of teaching strategies. The verbal protocols were separated into segments. When referring to *we* in this paper I refer to these four people. However I was the only one coding each segment and writing this paper.

To discover teachers' mental thought processes we address these following research questions:

1. Which mental thought processes are most frequently used while responding to the selected items?
2. Are any of the mental processes related to performance on the test?

3. Are there any differences between beginning students, advanced students and experienced teachers' mental thought processes while responding to the selected items?

1.1 Structure

Chapter 2 Theory and research: contains a theoretical framework for this research paper. It begins with a historical background on teachers' mental thought processes. Further it describes this process and teachers' decision making. In addition it provides prior research within the field.

Chapter 3 Methodological approaches: provides a description of verbal data. This was the methodological approach we used while asking teachers to think aloud. In addition it contains a discussion on this method's advantages and limitations.

Chapter 4 Material: describes the material used in our think aloud research study, the Arizona Educator Proficiency Assessment (AEPA). It also contains a description on licensure tests in general and a discussion on its validity.

Chapter 5 Method: explains the procedure of our study and describes the participants in our study. Further it contains a clarification of how we prepared the verbal protocol for analyses as provided by Chi (1997).

Chapter 6 Results: presents the research results. First descriptive statistics on cognitive processes are presented. Second, factor analyses with four new factors are introduced. Teacher level was used to measure differences. In the last part of this chapter regression analysis for variables predicting AEPA test score is presented.

Chapter 7 Discussion: discuss our results related to the theoretical framework provided in chapter 2.

Chapter 8 Conclusion: provides a summary of the most important findings.

2. Theory and research

2.1 Historical background

The studies on teachers' mental thought processes have their roots in the development of cognitive research in the late 50's and late 60's (e.g. Newell & Simon 1956). Research on teachers' pedagogical thoughts, judgments, decisions and behavior is developed out of research on human decision making (e.g. Shavelson, 1973, 1976; Shulman & Elstein, 1975) and problem solving (e.g., Shulman & Elstein, 1975). This suggests that its original orientation was psychological (National Institute of Education, 1975). Current research on teachers' mental thought processes concentrate essentially on three areas: studies of judgment and policy, studies of problem solving, and studies of decision making (Shulman, 1986). There are two fundamental assumptions research on teachers' thought processes relies on. The first assumption is that teachers are rational professionals who make decisions in a complex environment (e.g. National Institute of Education, 1975). The second assumption is that the human minds ability to create and solve complex problems is small compared to the size of some models of rationality (e.g., some normative model). Herbert Simon (1979) showed that individuals often have difficulty making economical rational choices, even though the decision scenarios were simple.

2.2 Teachers' mental thought processes

The need for research on teachers' intentions and thoughts and the relationship between them, and not just behavior alone, has several reasons. The main reason is that behavior alone cannot explain for predictable variations in teachers' behavior occurring from differences in their goal, judgments, and decisions. Teachers' thinking is unique to each individual. Teachers can have different perception of the same situation. Different beliefs and values can have a huge effect on the decision making. Such knowledge challenges the image of the teacher as a technician and the notion of research as a source of generalizable prescriptions. Consequently, research on teachers mental thought processes is important to provide an inside to the mental lives of teachers. However, in order to understand teachers' mental thought processes we need to consider the demands teachers experience.

2.3 Complexity of teaching

At the same time as teachers' are responsible for the planning decisions about what to teach, they are also in charge of all subject matter areas, how long to devote to each topic and how much practice to provide. This adds additional significance and complexity. Teachers often need to create their actions on the spot, knowledge about the immediate social environment is necessary to make these decisions. Prescription on how to teach spelling and grammar is of little practical significance if you do not take the contextual knowledge of learners into account. In addition, there is a great deal of uncertainty in the teachers' work, Clark and Lampert (1986) therefore define teacher knowledge as speculative. Thus we have to consider several factors involved in teaching, consequently the mental thought processes are affected. The decision making process is one of these mental thought process.

2.4 The decision making process

Teachers have to make decisions in the classroom all the time, consequently the decision making process is one of the most important mental thought processes teachers are involved in. Shavelson, (1973, p.18) describes decision making as follows: "Any teaching act is the result of a decision, whether conscious or unconscious, that the teacher makes after the complex cognitive processing of available information". This statement summarizes the most important aspect of teaching, and the reasoning leads to the hypothesis that the basic teaching skill is decision making.

However, is decision making a conscious choice? Decision has traditionally been implied as consideration of alternatives, nevertheless this has found little support within the research on interactive thoughts (Clark and Peterson, 1986). Clark and Peterson (1986) suggested that decision making during interactive teaching involved two or more alternative courses of action. If no alternative was available in the memory, the teaching would continue as before. Teaching is highly influenced by teachers' routine and intuitive teaching.

The decision making process consists of both planning and interactive decision making. The distinctions between teachers' interactive thoughts and decisions and their proactive thoughts

and decisions appear to be important. The thinking teachers do during interactive teaching does appear to be qualitatively different from the type of thinking they are engaged in when they are not interacting with students (Clark and Peterson, 1986). While interactive decision making occurs in interacting with students, teacher planning happens prior to classroom interaction.

In our study we selected items from Arizona Educator Proficiency Assessment (AEPA) practice tests while teachers were thinking aloud. The questions asked were about specific situations that may occur in a classroom. However, the teachers were not required to make the decision on the spot since there was no time limit. For that reason, the decisions teachers in our study made would probably be somewhere in between planning and interactive decision making. Thus I will consider both planning and interactive decision making when answering the three research question.

2.5 Which mental thought processes are most frequently used while teachers are to make decisions about how to teach?

Research on interactive decision making indicates that teachers encounter decision situations at two minutes intervals while teaching (Peterson and Clark, 1986), this tells us much about the positions teachers are in and the need to make on the spot decisions. However, there is a large difference between considered and not considered decision making. According to Oberg (2005) considered decision making processes only count for a small part of teachers' practice. In a large amount of teaching the circumstances are so complex that teachers act routinely or habitually, only a small proportion of teachers' practice is considered decision-making. Still, not everything happening in a classroom require decision making. According to Berlinder (2005) ordinariness and ongoing processes in the classroom do not require attention or the making of a decision. On the other hand you have atypicalness, irregularity and out of order events concerning students. These events require attention and some decision. The study on such events has developed a theory about the function of the perception. The theories indicate that perception is the primary mental process used during interactive teaching and students are the primary content of thinking during interactive teaching. These ideas origins came from information processing and cognitive psychology

(Berlinder, 2005). Also, previous studies indicate that selection is a complicated decision making process which involves the perception, assessment and evaluation of a variety of types of information (Bredeson. P, 1983, cited in Berlinder, 2005). The information you evaluate may possibly be correct, although it can also be inaccurate, irrelevant or even false. The evaluation used by teachers in both planning and interactive decision making normally concerns the students.

One of the earlier studies conducted by Taylor (1970) found that the most important subjects in teachers' planning were the pupils' needs, abilities and interests, followed by subject matter, goals and teaching methods. Zahorik (1975) found similar results in his study, decision mentioned by the greatest number of teachers concerned pupil activities, followed by content and learning objectives. Clark and Peterson (1986) also reviewed several studies describing teachers' interactive thoughts, all in which used stimulated recall interviews. Similar for all the studies was that the high proportion of teachers' thoughts was concerned with the learners.

Another perspective to understand teachers' thinking is to compare teachers' knowledge to the structures we use in everyday life, known as common sense. Some researchers argue the mental thought processes used to make decisions in the classroom are the same as used in everyday life. Buchmann (1987) argues the knowledge we use interacting with others on a daily basis is not much different from the one we use while teaching. She claims teaching is a natural part of human life. This means the common sense structures we use everyday obtain a huge position in teachers' thinking. Also teachers' own beliefs seem to have an effect on their mental thought processes and in turn action. Izu (1977) found that when relevant information was available for decisions, teachers used that information to make a decision (Cited in Clark and Peterson, 1986). But, when the information was not relevant for a decision, Shavelson et al. (1977) assumed that subjects' beliefs about education and teaching guided their decisions. Teachers' conception of a subject matter also is expected to influence their judgments, decisions and behavior (e.g., Barr, 1975).

2.6 Are there any coherence between teachers' mental thought processes and achievement?

Studies of teacher thinking can give us a description of the way teaching is and help us provide an educational framework on what sorts of information, advice and support will be useful in the classroom. To be a good teacher you need contextual knowledge, the context shapes teachers' thinking (Elbaz, 1983). Teachers often need to create their actions on the spot, understanding about the social environment is essential to make these decisions. In addition teachers need to take risks, the knowledge is transient rather than fixed and static (Clandinin and Connelly, 1986).

The relationship between both patterns of interactive decision making, teacher planning and student achievement has been examined within a few studies. The distinctions between teachers' interactive thoughts and decisions and their proactive thoughts and decisions appear to be important. The thinking teachers do during interactive teaching does appear to be qualitatively different from the type of thinking they are engaged in when they are not interacting with students (Clark and Peterson, 1986). Findings from research on teacher planning propose that there are several reasons to why teachers plan the way they do. Clark and Yinger (1979) clustered teachers' responses about why they plan into three categories: 1. planning to meet immediate personal needs (e.g. reduce anxiety), 2. planning as a resource to the end of an instruction (e.g. to learn the material and to organize time), 3. planning to serve a direct function during instruction (e.g. to get an activity started). The literature is almost exclusively descriptive and deals primarily with the planning of experienced teachers which specify that we do not know how experienced teachers planning differ from that of novices (Clark and Peterson, 1986). However teacher planning seems to be a major determination of what is taught in schools.

Research on interactive decision making has pointed towards that high achievement test scores is characterized by rapid judgment, reducing many events and cues into a small amount of categories and willingness to change the course when needed (Clark and Peterson, 1986).

This will also be discussed in the next section, given that expert teachers often are associated with high achievement.

2.7 Are there any differences between novice and experienced teachers' mental thought process?

The sample in our study consisted of beginning students, advanced students and experienced teachers. Although most of the research existing in contrasting group studies on mental thought processes consisted of novices versus expert teachers. Hence, I will focus on these studies. The experienced teachers in our study are closer related to the expert teachers, while the beginning students can be compared to the novices. Advanced students are somewhere in between the novice and expert teacher.

Novice – expert contrasts have provided information in a numerous of areas e.g. in the study of complex cognitive tasks such as playing chess (Chase and Simon, 1973), solving physics problems (Champagne, Gunston and Klopfer, 1983; Chi, Feltovich and Glaser, 1981) and note taking (Heidi and Klainman, 1983). Experts are usually associated with experience and high achievements. Novices are typically described as someone new to the field.

The purpose of expert novice studies is to find differences between beginning and qualified professional and to expose the interval stages in the development of expertise. This can help us promoting the achievement of expertise. The goal is to find out what kind of characteristics are representative for effective teachers.

Contrasting groups are frequently used in studying teacher thinking. Berliner (2005) is among the researchers who recommends designing more studies using contrasting groups. If the teachers in our study are of contrasting groups we usually learn much more than if they are not. One approach is to find expert and novice teachers, or like we did in our study, where we studied the differences between beginning students, advanced students and experienced students. According to Berliner (2005) contrasting groups are already used in about 10 percent of the studies done in teacher thinking, e.g. Clark and Peterson, 1986. Also the effective schools literature (Ron Edmonds, 1979 and Rutter, 1979), emphasise and often

use the method of contrasting groups. Berliner (2005) states that research on teacher thinking can benefit from this contrasting groups approach. A number of sources and studies point towards a difference in how novice and experienced teachers perceive atypical events. Atypicalness, irregularity and out of order events concerning students require attention and decisions. Experienced teachers appear to handle these events better than less experienced teachers (e.g. Berlinder and Carter, 1986).

In a study of contrasting novice and expert competence in math lessons, Leinhardt (2005) compared the performance of experienced and highly competent or expert teachers with new or novice teachers. They found expert teachers to have richer and more detailed plans for their lessons. There was also a difference in the two groups' agenda for the class. Testing students in terms of checking students understanding which would influence teachers' decision on whether or not to go on to the next instructional move was hardly mentioned by the novice teachers in the interview about their agenda. Experts' agendas were richer, more detailed, had goal statement and actions. Another difference was that all experts in their study started the lesson with telling what they had done the day before and saw lessons as tied together as a pattern. Expert teachers gave better explanations than the novices did and they used the same representation for multiple explanations. An additional distinction between novice and expert teacher had to do with subject matter. Novices did not seem to know the subject matter well enough to be flexible while teaching it.

The results of a study conducted by Leinhardt and Greeno (1986) demonstrated that expert teachers constructed their mathematic lessons around a core of activities, which seemed to move from total teacher control to independent student work. The routines during those sessions were more efficient by experts. Novices often jumped directly from presentation to practice. There was a major difference when it came to routines, experts used more flexibility. Greeno (1984) found experienced teachers to be distinguished by their ability to maintain control of their agenda at the same time as they retained new information in interaction with students. Erickson (1984) found the same in his study about simultaneity, focusing on how teachers make sense of what happens in their classrooms. He found that experienced teachers appear to know what to look at and what to ignore. Experienced

teachers hardly ever noticed things in isolation, and concentrated mostly on situations and events that called for some decision and action. In contrast, teacher students focused on the most uncontrollable students. As a result, teacher students appeared to be more worried with management than with instructional strategies or students learning. Keeping order was the primary concern.

There is also a difference in how to apply their knowledge. An expert knows better how to apply their knowledge. A novice may possibly have enough knowledge about the problem situation, but does not know how to apply this knowledge (Glaser, 1985). In addition there appears to be a difference in metacognitive capabilities. These capabilities include knowing what one knows and doesn't know, planning ahead, monitoring and one's effort to solve a problem.

To summarize we might say that expert teachers seem to be related to the structure and quality of the teacher's knowledge base, which is both declarative and procedural. Experts are able to monitor, understand, and interpret events in more detail and with more insight than either novices or advanced beginners. They seem to differ in their abilities to monitor and interpret simultaneous events within a classroom. The relation between the knowledge base and problem solving seems to be determined by the quality of the representation of the problem. Experts have much better representation of the problem and can base their solution from such representations.

2.8 Expectation

After reviewing the literature what would one expect to find in our study when trying to explore our research questions?

1. Which mental thought processes are most frequently used while responding to the selected items?
2. Are any of the mental processes related to performance on the test?
3. Are there any differences between beginning students, advanced students and experienced teachers' mental thought process while responding to the selected items?

From reviewing the literature on teachers' mental thought processes we have learned that in a large amount of teaching the circumstances are so complex that teachers act routinely or as a habit. Only a small proportion of teachers' practise are considered decision-making. The complexity may explain the high proportion of thoughts directed toward learners. Teachers' concerns about content, materials, objectives might fade or are redirected into concerns about student's learning. The students provide the major source of cues that gives rise to interactive decisions. The major proportions of teachers' thoughts are about students, followed by instructional behaviour and procedures, content materials, and learning objects. As a result, one does not expect to find a lot of considered decision making processes in our study. According to Oberg (2005) these only count for a small part of teachers' practise. Possibly the results would consist more of mental thought processes based on routines, habits or common sense. One could also suspect that teachers' beliefs about education and teaching guided their decisions.

Research on mental thought processes and achievement has provided information about the need for contextual knowledge. To create their decision on the spot, knowledge about the immediate social environment is necessary. Also planning seems to be important, however research have not yet fully discovered what actually distinguishes high achievers from low achievers. Also high achievement test scores are characterized by rapid judgment, reducing many events and cues into a small amount of categories and willingness to change the course when needed. However we used Arizona Educator Proficiency Assessment, which is a multiple choice test to discover teachers' mental processes. It might be difficult to discover their contextual knowledge and their willingness to change course when needed.

One assumes there would be a difference between beginning students, advanced students and experienced teachers. Experienced teachers seem to be related to the structure and quality of the teacher's knowledge base, which is both declarative and procedural. Experts are able to monitor, understand, and interpret events in more detail and with more insight than either novices or advanced beginners. They seem to differ in their abilities to monitor and interpret simultaneous events within a classroom. The relation between the knowledge base and

problem solving seems to be determined by the quality of the representation of the problem. Experienced teachers have much better representation of the problem and can base their solution from such representations. One presumption is that differences in representation of the problem come from more experience. Consequently, one predicts experienced teachers draw more on experience in answering the items in our study.

3. Methodological approaches

An important difference between teachers' thought process and teachers' action is the extent to which the processes involved are observable. Teachers' thought process take place "inside teachers' heads" and as a result it is not observable. Teachers' behaviour is in contrast an observable phenomenon and consequently more easily measured and by far more subjected to empirical research. Research on teachers' thought processes depend heavily on a variety of self-report forms. As a result teachers' thought processes present several challenging methodological problems including how to elicit and interpret valid and reliable self-report about cognitive processes. We also have to take context into account and not concentrate on the teacher alone, Huber (2005) argues teacher thinking should always include the whole field in which a teacher's activities are surrounded by.

Clark & Peterson (1986) reviewed several studies on teacher thinking, and researchers frequently used various combinations of five methods: thinking aloud, stimulated recall, policy capturing, journal keeping and the repertory grid technique. These methods are often supplemented by interview, observation and case studies, however these methods can also be used as the main research method in the study of teacher thinking.

3.1 The purpose of the study

The purpose of this study was an attempt to understand the mental thought processes used while taking tests similar to the practise AEP items used in our study. To investigate this, we studied the cognitive process of 45 beginning students, advanced students and experienced students while they were thinking aloud. We used verbal report analyzes to collect evidence. This method has also been referred to as think aloud study or the method of verbal protocol analysis (Ericsson, 1987). This technique is a major method in cognitive behaviour research. Verbal protocols analysis has increasingly been used to support different educational measures (Leighton, 2005).

To become a licensed teacher in the United States, you have to complete and pass a test similar to the items used in our study. According to Arizona Educator Proficiency Assessment, it's expected to measure pedagogical knowledge and skills. Further insight into response processes can help us understand the underlying mental process in use while responding to items similar to the ones used in our study.

In addition we wanted to explore if these tests in fact are measuring pedagogical knowledge and skills and if there are differences between the three groups' mental thought processes.

3.2 Analyses of verbal data

One technique of collecting support about the fundamental factors underlying test scores is to examine the substantive cognitive process that drive examinee response behavior (Messic, 1989, Cronbach, 1971). This approach tries to explain mental thought processes of participants when they responds to test items similar to the ones in our study, and decide whether those processes are consistent with those predicted by the theory. Cognitive processes can be compared to explicit cognitive models or more general assumptions about task performance to validate a given test. Chi (1997) defines verbal analysis as: "verbal analysis is a methodology for quantifying the subjective or qualitative coding of the contents of verbal utterances" (p.273).

The verbal analysis is well-established in research that tries to understand cognition. The aim of this technique is to capture the representation of knowledge that a learner has and how that representation changes with achievement. Using this method can help us understand questions like what separate the knowledge of an expert or a more advanced learner with that of a novice. It has been used for example to code explanations of what one understands as one read a text, to see whether an explanations is an inference, a monitoring statement or some other irrelevant comment. Further it tries to understand what a learner knows and how that knowledge influences the way the learner reasons and solves problems, whether correctly or incorrectly. The standard method for getting subjects to verbalize their thoughts concurrently is to instruct them to "think aloud".

Ericsson and Simon (1993), who have been important in the development of this approach, suggest that while students are performing a think aloud study, this approach is capable of capture what's held in the short time memory. This results in a sequence of thoughts that reflects what occurs cognitively during completion of a given activity.

The most common methods used in educational research are observation while working, examination of work samples, evaluation of assessment data and questioning of students their teachers and caregivers. All these methods enhance our understanding on how people learn. However if you get students to articulate their thoughts while engaged in an activity it can help us understand the underlying mental thought processes in use while engaged in a given activity. Capturing what they are verbalizing provides extensive information which cannot be provided with any of the other methods. Analyzes of verbal data can therefore be a valuable method in gathering further knowledge on peoples' thought processes.

Capturing think aloud data requires the participants to continually speak what they are thinking as they work. These thoughts can be collected at various points, but it's generally collected when the participant is in the processing of gathering information, which is referred to as concurrent probing. This was how the participants thought processes were collected in our study. However according to Ericsson & Simon (1993) they can also be captured after completion of the given activity, which is referred to as retrospective probing.

Payne (1994) claims that some tasks are better suited to be studied using verbal protocol tasks, he further says that the more a task involves a cognitive process that take more than a few seconds to perform and the more the task involves verbal types of information the more suited they are to think aloud study. Branch (1994) proposes that problems could possibly occur in eliciting verbal data if:

- the task involves a high cognitive load

- the information is difficult to verbalize because of its form.
- the process is automatic for the participant.

Occasionally it can be difficult to presume if the task involves a too high cognitive load or if the process is automatic for the participants, everything depends on the participants involved in your study. After analyzing our data in this study, we can assume that the processes were not automatic for neither of the participants. The overall score in the test also indicates that the task did not involve a too high cognitive load. Further the form used point towards that it was not difficult to verbalize as most of the participants were highly verbal.

Analyses of verbal data can take several forms, such as those already mentioned attained through concurrent or retrospective probes, but they can also emerge through more naturalistic situations like interviews and peer interactions.

Studies of analysis of verbal data have traditionally relied on participants' verbal reports of their cognitive processes as they complete assessments task. Bloom and Broder compared think-aloud protocols of high and low scoring examinees on a test of reasoning in a 1950 study, (Bloom & Broder, 1950). More recently, think aloud protocols have also been used to help validate online problem solving assessments (Chung et al., 2002) and to help collect evidence about the validity of scores from tests in the National Education Longitudinal Study of 1998 (Hamilton et al., 1997).

3.3 Advantages of analysis of verbal data

One of this method's major advantage is that it reduces problems associated with memory failure which may occur when one waits to collect verbal data at the conclusion of an activity (Wade 1990). Participants' capacity to remember what they were thinking at an earlier time is often limited, even with stimulus materials like video recall. Further this method is considered more reliable because the participant is engaged in a real activity and not requested to make decisions on a hypothetical situation (Wade 1990). Asking an

individual how they would complete an activity, without actually engaging in the activity, would make accurate verbal data far more difficult to produce.

Ericsson & Simon (1993) state that concurrent verbal protocols can offer a “dramatic increase in the amount of behaviour that can be observed when a subject is performing a task while thinking aloud compared to the same subject working under silent conditions” (p.13).

Analysis of verbal data can also be used in addition to observational data which is a common method used in qualitative research. One of the problems with observational data is the interpretation of it. William & Clarke (2002) looked at this issue in a study. Their findings indicated that if you added the student’s voice to the study it provided formulation of ideas, further evidence of thinking and a higher level of selective attention. This study points towards the neglect of some of the most helpful supplies in understanding learning in a lot of instances.

3.4 Limitations of analysis of verbal data

Problems with drawing conclusions from verbal protocols can take place because of participants’ reactivity or veridicality (Russo, Johnson, & Stephens, 1989). Reactivity occurs when verbalization changes the main mental process, by disturbing the nature or speed of participants’ responses. Veridicality refers to the degree which the participants are able to report their own thinking precisely without forgetting relevant information or reporting thoughts that have never taken place.

There are three main effects you need think about when you refer to participants’ reactivity. The first refers to the participants’ ability to think and taking a task at the same time. The second concern refers to the effect of talking aloud under circumstances which usually would take place in silence. The third refers to the effect of participants’ attention of the cognitive process underlying while attaining a task.

Some researchers have claimed that we have a limited capacity to work through an activity and talk at the same time (Stratman & Hamp-Lyons 1994, cited in Branch 2000). As an effect of this, a few individuals can have a reduced capacity to talk at the same time as they are working on a task. This can sometimes be seen when the participant turns silence during a task. In addition we have to consider the problem related to verbalizing under circumstances which is usually undertaken silently. However if you try to maximise the participants' comfort while engaging in a task, you can be able to reduce this effect. In our study we had one practice question which can be helpful to get the participant used to the situation.

There has also been raised concern about gaining depth in the data produced. Branch (2000) reported that while transcribing 144 verbal analysis protocols she observed that a number of the participants had very short or procedural protocols. A problem can be that the verbal analysis protocols do not reflect the participants' abilities, which in fact can underestimate the participants' learning abilities.

In spite of all this Ericsson and Simon (1993) reviewed several experimental and quasi experimental studies, and concluded with performance was not affected by verbal processes when participants verbalized simultaneously with task performance without introspection. Also Nisbett and Willson (1977) documented that verbal reports could be correct in some contexts, in particular when you obtain response and stimulus close in time. They state that "reports will be accurate when influential stimuli are available and plausible causes of the response, and when few or no plausible but noninfluential factors are available" (p.253).

Another issue when it comes to verbal analyze has to do with analyzing the data from the verbal protocols. These methods can be classified into three groups:

1. Quantitative/exploratory

Chi (1997) describes this approach as "the qualitative data is examined for impression and trends, methods of coding are developed to capture those impressions, and codings can then

be analyzed quantitatively” (p.7). Several analyses can be performed within this approach, for example relating mental thought strategy to performance or comparing differences in reasoning between novices and experts. This was the approach used in our study

2. Quantitative /confirmatory

This approach obtains models for problem-solving from theory and task analysis. The matching of verbal protocols with the theory based-solving model is used as a measure of the accurateness or competence of the protocols. Similarity between models and protocols are assessed by judgmental methods or computer algorithm. In a lot of cases the construction of a valid computer model is the main goal of this type of research (Ericsson& Simon, 1993).

3. Qualitative

This approach tries to discover and describe the underlying structure the cognitive process in use through the use of reason-based judgment. The goal is to articulate generalities across individuals (Aanstoos, 1982).

Even though there are some limitations regarding participants’ reactivity or veridicality, and problems regarding analyzing the data from the verbal protocols, I still believe the strengths outweigh the limitations. Several studies including a few discussed in this paper have been given a further understanding on participants’ thoughts and ideas. I believe this is the best technique to understand these underlying mental processes participants use while answering questions similar to the items used in our study and further to answer our research questions:

1. Which mental thought processes are most frequently used while responding to the selected items?
2. Are any of the mental processes related to performance on the test?
3. Are there any differences between beginning students, advanced students and experienced teachers’ mental thought process while responding to the selected items?

4. Material

The test items used in our study were taken from Arizona Educator Proficiency Assessment (AEPA). We used two different forms in this study. Form A consisted of eight question, all which were followed by four multiple choice items. Twenty-five participants responded to form A. Form B consisted of nine questions with the same structure as form A. Twenty of the participants responded to form B. Three of the questions were present in both form A and form B. Each of these questions presents an educational situation or challenge with four possible explanations (See appendix). Participants were instructed to verbalize to all the alternatives, also the ones they didn't choose.

The test we used are supposed to measure basic skills, general knowledge, content knowledge, and knowledge of teaching strategies. They are used to separate at least minimally competent examinees in the areas assessed from those who are not. However we wanted to discover if these licensure tests in reality separated competent from incompetent beginning teachers.

4.1 Licensure tests

Americans have taken on a reform programme for their schools that calls for excellence in teaching and learning. School officials across the nation are aspiring for instruction at high levels for all students. However there are still gaps between the nation's educational ambitions and student achievement. As an attempt to reduce these gaps, policy makers have recently focused on the qualifications of teachers and the preparation of teacher candidates. One of the effects of this ambition has increasingly been that states are testing candidates who want to become teachers. 42 states require candidates to pass one or more tests of basic skills, general knowledge, subject matter knowledge, or teaching knowledge (National Associations of State Directors of Teacher Education and Certification, 2000b, cited by National Research Council, 2001, p. 44). Teacher licensing relies on more than tests to judge whether candidates have the knowledge, skills, abilities, and dispositions to practise responsibility. Usually they must fulfil education requirements, successfully complete practise teaching, and receive the recommendations of their preparing institutions. In states

where tests are required, candidates can leave without a degree because they fail the licensing test, even though they successfully have completed all of the institutional requirements for graduation. The candidates who fail this test will also lack a licence to teach in the state's public school system. However, teacher candidates who fail the test can and do teach in many private schools in some states. In fact, large numbers of individuals teach with emergency licenses (National Research council, 2001).

4.2 Arizona Educator Proficiency Assessment (AEPA)

The test items used in our study were taken from Arizona Educator Proficiency Assessment (AEPA). The test items were drawn from practise AEPA test, which have been used in expired AEPA tests. However they were currently only used as practise items.

AEPA is a professional knowledge test, which is required of all applicants for a teaching certificate in Arizona. There are three levels of the Professional Knowledge test: Early Childhood, Elementary, and Secondary (www.aepa.nesinc.com). Everyone applying for a certification to become a teacher is required to pass an examination. The test Early Childhood, Elementary, and Secondary are suppose to measure pedagogical knowledge and skills. The test consists of approximately 100 multiple-choice questions like the ones used in our study, and three written performance assignments. The purpose of these tests is to identify candidates who have demonstrated the level of professional knowledge and skills judged to be important for Arizona educators (www.aepa.nesinc.com).

4.3 Concerns regarding licensure tests

A number of questions have been raised about whether current tests measure what they are supposed to measure. One of the criticisms against these tests is that they rely on multiple choice or short answer questions to short statements of professional problems. This may fail to represent the complexity of the decision-making process. This was also exemplified in our study where several of the participants talked about the complexity of the given situation and argued answering this question would be difficult without knowing more about the situation. A few assumed that more than one alternative could be the right answer; while others argued

that none of the alternatives would be the right thing to do. The following two examples from the study illustrate this problem:

“...we don’t know anything about how her dad treats to her at home. The only thing she says is: I know it’s wrong. Dad says writing is hard for me...”

“...actually I would not choose any of the alternatives...”

A lot of the questions are also likely to depend on different factors. A thoughtful teacher making decisions would need to know many things not treated in the questions. These decisions can be difficult to capture amongst these tests. Since a teacher’s work is complex, still a well-designed test would have problems measuring all the fundamentals competencies a teacher ought to know. These tests can simply offer some information essential to evaluate the competencies of teacher candidates.

4.4 Validity evidence of licensure tests

The standards for Educational and Psychological Testing 1999 standards say that “validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests” (American Educational Research Association, et al., 1999:9 cited by National Research Council, 2001). The primary purpose of licensure testing is “to ensure that those licensed possess knowledge and skills in sufficient degree to perform important occupational activities safely and effectively” (American Educational Research Association, et al., 1999:9, cited by National Research Council, 2001).

Validity research on licensure tests focuses “mainly on content-related evidence, often in the form of judgments that the test adequately represents the content domain of the occupation” (National Research Council, 2001, p.157). In general validity evidence for employment require a clear definition of the occupation, a clear definition of the nature and requirements

of the job and expert judgment on the fit between test content and the job's requirement. This is essential to understand the role of licensure tests, however this can be difficult because the criteria for doing so vary from person to person, from one community to another, and from one area to the next. What considered teacher quality has changed over time as society has shifts its values. Further it takes knowledge, skills, abilities, and dispositions needed to perform job duties and tasks to an account. The profession is expected to have data to demonstrate the relationships between test results and criterion of interest. However in some cases this is difficult to obtain.

A number of researchers have explained the measurement and design difficulties associated with gathering job-related information for beginning teachers. Measuring distinguishes between minimally competent and minimally incompetent beginning practise can be problematic, in particular because candidates are working in many different settings. Smith and Hambleton (1990) claim using student achievement data as a criterion measure for teachers' competence can be difficult because you measure and isolate students' prior learning from the effects of current teaching. Some isolate school and family resources that interact with teaching and learning. In addition it is difficult to follow teachers and students over time and take multiple measurements due to its time and resource matter.

It has been claimed that more content-related evidence is essential for establishing the validity of teacher licensure tests. The committee on Assessment and Teacher Quality has been calling for empirical evidence on the relationships between performance on teacher licensing tests and other relevant variables. The 1999 standards suggests gathering additional validity evidence in analyzing mental thought processes in which examinees engage in as they respond to assessment exercises and assessments components (National Research Council, 2001). This is what we are trying to get a further understanding of in our study. The standard also suggests that you can gather additional validity if you examine patterns of relationships among assessment exercises and assessments components, and correlate measure of the same and different constructs.

Paggio and colleagues (1986) gathered helpful data on the relationship between licensure test results and other measures of candidate knowledge. They gained evidence of validity by comparing the performance of education and noneducation majors at the University of Kansas on one of the precursor test to Praxis-the National Teachers Examination Test of Professional Knowledge (National Research Council, 2001).

There has been a disagreement in the field about the type of validity evidence that should be collected for teacher licensure tests. Jaeger (1999) is among the ones who claim criterion validity virtually impossible to obtain (cited by National Research Council, 2001). This reflects the committees' framework for evaluating teacher licensure tests. This framework does not essentially call for validity studies that examine the relationship between performance on the tests and future performance in the classroom. The committee does consider whether empirical evidence has been collected on the relationships between performance on licensure exams and other concurrent measures of knowledge and skills similar to those covered on the exams.

5. Method

5.1 Participants

The sample consisted of forty-five participants with a variety from beginning students to experienced teachers. The students were recruited from University of Arizona from the department of Teaching and Teacher Education. The teachers were recruited from local Tucson public schools. The participants got twenty-five dollars each to participate in the study.

The sample was composed of thirty-six Caucasians, four had Latinos origin, two were Asians and three were from other origins or would not respond. Four among the sample was male. All of them were Native American speakers. The participants differed when it came to education and experience and were classified into three different categories:

1. Beginning students

Twelve of the participants belong to this category. The students in this group were beginning students which had been taking fundamental educational courses and had no teaching experience

2. Advanced students

Eighteen of the participants belong to this category. These were advanced students in their final year of studying. The students had completed or were currently doing their internship. All of them had experience from student teaching.

3. Experienced teachers

Fifteen of the participants belong to this category. The participants in this category were experienced teachers with a range from 1 to 28 years of experience, with a mean of 15.27 years of practice and std. deviation of 8.99.

Table 1: Teaching experience

| Years of experience | Frequency | Percent |
|---------------------|-----------|---------|
| 0 | 18 | 40.0 |
| 1 | 1 | 2.2 |
| 4 | 1 | 2.2 |
| 5 | 1 | 2.2 |
| 9 | 1 | 2.2 |
| 10 | 1 | 2.2 |
| 11 | 1 | 2.2 |
| 14 | 3 | 6.7 |
| 17 | 1 | 2.2 |
| 22 | 1 | 2.2 |
| 25 | 1 | 2.2 |
| 27 | 1 | 2.2 |
| 28 | 2 | 4.4 |
| Student teaching | 12 | 26.7 |
| Total | 45 | 100.0 |

The entire teacher group had already passed the AEPA test. Eleven of the students had taken the test and fourteen had not taken it or would not respond. Four of the students had passed it. Twenty-one had not been taken it, would not respond or failed the test.

5.2 Procedure

The items used in our study were administered to each participant individually. Participants were asked to think aloud about each of the four alternatives about their responses while being audiotaped, following a standard protocol that met the requirements for current verbalization as described by Ericsson and Simon (1993).

There were four different interviewers who interviewed approximately eleven participants each. The participants were asked to make mental inferences about each of the alternative to every item, also the ones they didn't choose. There was no time limit on the items, in contrast to the AEPA test which has a time limit. On average the test lasted for approximately two hours.

The participants were instructed to reflect all their thoughts without editing their earlier opinion. Participants were reassured that all thoughts were of interest for the study. There was one warm up question, once warm up was done, the interviewer tried not to interfere, except when they had to clarify what the participant said, or remind the participants to continue verbalizing an alternative.

5.3 Preparing the verbal reports for analysis

Once verbal data was collected and transcribed, we analyzed the entire protocol. Following the administration of the verbalization items was transformed into transcripts. After transcription, the verbal protocols were separated into segments. Each segment was coded independently. Different approaches are provided by Chi (1997).

Complete verbal protocols may be divided at various points, "revealing units of varying grain sizes, such as proportion, a sentence, an idea, a reasoning chain, a paragraph, an interchange as a conversational dialogue, or an episode" (p.9). Segments were defined as meaningful utterances separated by syntax of grammatical subordination or long pauses.

Two examples of a segment protocol are given below:

...she may need a more instructional approach to writing//

but she is only in first grade//

most first graders have trouble completing a sentence//

her construction of sentence is fine...

...it doesn't look like he did very well on the assignments//

and I don't know if there was any variation in the quality of instruction//

and there might be plenty of challenge in topic 4...

This approach to coding the segmented items responses was based on the quantitative/exploratory method described for quantifying data provided by Chi (1997). The purpose of this coding strategy was to categorize the segments into types of processes that could be used to address the research question.

In this method the analytic processes proceed from a set of categories used in the present study were: Repeating the Question, Repeating the Alternative, Common Sense, Learning Theories-based Inferencing, Research-based Inferencing, Rehearsal Cues, Elimination, Do not know the Concept, Drawing on Own or Others Experience, Deciding about Option and Nonsolution-Productive Thinking. When the preliminary categories were found not to cover the protocols adequately, additional codes were created or original codes modified, until a coding system was developed to cover the protocols adequately. Chi (1997) calls this "piloting the analysis" (p.8).

Once a sufficient system for coding was developed, the rest of the segmented transcript was coded into different categories.

The final system for coding contained 19 categories. Each of them is briefly described and followed by an example in the following table:

Table 2: System for coding

| | Category | Description | Example |
|---|---------------------------|---|--|
| 0 | Uncodable Items | Segment which was difficult to understand or impossible to code belongs to this category. | "...mumbling" |
| 1 | Repeating the Question | Segment which repeated the question either directly or changed the wordings of it. | "...A teacher whose student achieve at a range of levels is considering using ability grouping for much of the instructions in his classroom. According to educational research, such an approach is most likely to have which of the following effects?..." |
| 2 | Repeating the Alternative | Segments which repeated the alternative either directly or changed the wordings of it. | "... Increasing obstacles to achieving a sense of community and mutual support among all students..." |

| | | | |
|---|-------------------------------------|---|--|
| 3 | Common Sense | A form of evidence that is based on common knowledge, conventional wisdom, tradition, or someone's personal philosophy or perspective belongs to this category. | <p>“...You have to make use of transition times...”</p> <p>“...otherwise time just slips away and gets eaten up...”</p> |
| 4 | Learning Theories-based Inferencing | Segments which applicants used principles from learning theories to solve the problem. | “...received instead of used, that goes with the progressive theory of education that children are not just empty vessels to be filled...” |
| 5 | Research-based Inferencing | Segments which applicants used principles from research to solve the problem. | “...current research suggests that this would have an effect on reading...” |
| 6 | Rehearsal Cues | Segments in which participants paraphrased the fact situation, paraphrased elements of fact provided in the response option, referenced elements of given facts without interpretation or inference, or reinforced memory of facts by making the test form belong to this category. | “...It could be working, but probably not...” |
| 7 | Elimination | Segments which used deductive elimination strategies belong to this category. | “...Number 2 is a stronger negative effect than number 1...” |

| | | | |
|----|---|---|---|
| 8 | Do not know the Concept | Segments were the participant did not understand the concept which appeared in either the question or alternative belongs to this category. | "...I don't even know what meta-cognitive approach to reading is..." |
| 9 | Do not know | Segments were the participants did not know the answer. | "...I just don't know..." |
| 10 | The question have nothing to do with the Question | Segment in which participants referred to the alternative as irrelevant to the question. | I don't believe that A is the right answer because he didn't ask the teacher for assistance" |
| 11 | Mental Error | Segments which not correlate with the conclusion the participants made. | "...This is defiantly unlikely..." "...I'm going to say A which is likely..." |
| 12 | Own or Others Experience | Segments were participants draw on own or others experience. | "...The reason why I'm choosing this is because as a teacher you see how far removed parents can be..." |
| 13 | Learned in Courses | Segments were the participants use what they've learned in courses or read belongs to this category. | "...I've learned in class that you should plan, teach, and then carry out what you're teaching..." |
| 14 | Visualizing | Segments were participants describe what they would do if they were in the situation explained in the question. | "...I know if it was me as a student, I would think, do I deserve that?" |

| | | | |
|----|---------------------------------|---|--|
| 15 | Deciding about Option | Segments s in which participants expressed decisions about choices. | "...I just don't think that sounds right..." |
| 16 | Nonsolution-Productive Thinking | Segments were participants verbalized in ways that did not advance their problem-solution, for example task irrelevant information. | "...Yeah I didn't give that much thought did I?..." |
| 17 | Explaining the Concept | Segments were participants explain the meaning of a concept. | "...well a meta-cognitive approach to reading is how our kids thinking when it comes to reading, what goes through their minds..." |
| 18 | Not enough Information | Segments in which participants expressed concern about the lack of information provided in the question belong to this category. | "...there's not enough information to answer this question..." |

6. Results

We first examined the proportion of each mental thought process. The statistics are based on the number of segments belonging to each category in each item response for each person, divided by the total number of segments belonging to each category in each item response for each person. As a consequence participants who used more mental process segments got a lower mean average of for example Learning theory-based inferencing than someone who used the equal number of the same mental thought process, but a small amount of coded segments. There was a large variability among length of the verbal responses among the participants. The number of segments in the verbal response ranged from six to sixty-five, and appeared to be an effect of individual differences rather than differences between items. This will affect the overall mean proportion. Score based entirely on frequency of statements would give a systematic bias to participants who have a tendency to express large quantities of ideas, including largely irrelevant thoughts.

Deciding about the option was the most frequent mental thought process across the segmented items, almost a quarter of all the mental processes were coded into this category. A typical segment classified as deciding about options would be “...I don’t think alternative A could be right...” Deciding about the option was followed by Rehearsal Cues, Repeating the Option and Common Sense as the most frequent thought processes. These four categories can count for almost seventy percent of all the segments. The standard deviation of these averages indicates that there was a substantial variability in proportion of segments in the major categories. The smallest proportion of mental thought processes were classified into Not Enough Information, followed by Do Not Know the Concept, Mental Error, Learned in Courses, Learning Theories-based Inference, Research-based Inference, Explaining the Concept and Do Not Know. These categories were seldom or never used for a major proportion of the participants in our study. The results suggest that participants did have difficulty explaining why they made their decisions. If we look at three of the four most common mental thought processes, deciding about options repeating the alternative, and rehearsal cues. A typical series of segments exemplifying these mental thought process could be like this: “....well, alternative A seems like the right answer...”, “....present a portfolio or other examples of Bethany’s work and point out specific ways in which she is making

progress...”, that could be the right answer, but I have to think this through first...”. This mental thought process of segments does not rely on pedagogical knowledge and skills. On the other hand, mental thought processes which rely on pedagogical skills and knowledge, found in the categories: Learning Theories-based Inference, Research-based Inference,

Explaining the Concept and Learned in Courses can only be found in just above two percent of the total segments.

Table 3: Descriptive statistic on cognitive processes

| Mental thought process | Mean proportion | Standard deviation |
|--|-----------------|--------------------|
| Not enough information | .001 | .00 |
| Do not know the concept | .003 | .01 |
| Mental error | .004 | .02 |
| Learned in courses | .004 | .01 |
| Learning theories-based inference | .005 | .01 |
| Research-based inference | .006 | .01 |
| Explaining the concept | .008 | .01 |
| Do not know | .012 | .01 |
| Visualizing | .015 | .02 |
| Own or others experience | .021 | .03 |
| Non-solution-productive thinking | .037 | .03 |
| Repeating the question | .055 | .04 |
| Elimination | .064 | .06 |
| The question has nothing to do with the question | .075 | .03 |
| Common sense | .103 | .06 |
| Repeating the alternative | .166 | .09 |
| Rehearsal cues | .176 | .07 |
| Deciding about option | .239 | .07 |

After analyzing these data we subsequently conducted a factor analysis to reduce correlational data to a smaller number of factors as an attempt to create a composite score. We created a correlation matrix, with our classifications for the mental thought processes as a basic variable that accounted for the interrelations observed in the data.

Table 4: Component Matrix

| | Component | Component | Component | Component |
|-----------------|-----------|-----------|-----------|-----------|
| | 1 | 2 | 3 | 4 |
| Average code 0 | -.483 | -.444 | -.126 | -.111 |
| Average code 1 | -.572 | -.291 | .117 | .177 |
| Average code 2 | .122 | -.117 | .327 | .334 |
| Average code 3 | -.234 | .820 | -.082 | -.227 |
| Average code 4 | -.131 | .145 | .162 | .514 |
| Average code 5 | .618 | .686 | .152 | .111 |
| Average code 6 | .019 | .051 | -.040 | -.375 |
| Average code 7 | .019 | .282 | -.457 | .290 |
| Average code 8 | .643 | -.161 | .178 | .356 |
| Average code 9 | .761 | -.138 | .039 | .210 |
| Average code 10 | -.122 | .160 | -.759 | -.145 |
| Average code 11 | .528 | -.462 | -.006 | .081 |

| | | | | | |
|---------------|------|-------|-------|-------|-------|
| Average 12 | code | -.108 | .225 | .650 | -.257 |
| Average 13 | code | .053 | .174 | -.218 | .486 |
| Average 14 | code | .121 | .184 | .612 | -.228 |
| Average 15 | code | -.155 | -.641 | -.085 | -.265 |
| Average 16 | code | .687 | -.137 | .068 | .071 |
| Average 17 | code | -.004 | .001 | .139 | .697 |
| Average 18 | code | -.139 | -.106 | .251 | -.149 |

We eliminated five of the mental thought process classifications: Uncodable, Elimination, The Answer have nothing to do with the Question, Learned in Courses and Not Enough Information as they did not have a relationship to any of the other categories, were rarely used, meaningless or insignificant to the study.

A number of the variables were correlated with each other and seemed to emerge at the same time, based on these results we created four new factors:

1. Common sense

Common sense and deciding about option belonged to this category. These two categories were related, but were negative correlated. Consequently when they didn't rely on common sense they had a tendency to rely on deciding about option.

2. Theories

The mental thought processes which belong to this group are Learning Theories-based Inferences, Research-based Inferences and Explaining the Concept.

3. Own Experience

Own or Others Experience and Visualizing were the two mental thought processes which were classified into this category.

4. Error

The fourth group consisted of these following mental thought processes: Repeating the Question, Repeating the Alternative, Rehearsal Cues, Do not Know the Concept, Do not Know, Mental Error and Non-Solution Productive Thinking.

We used teacher level as a factor to measure differences between these four groups and teaching levels.

Table 5: Summary results of teaching level as a component

| | Teaching level | Mean | Std. |
|----------------|----------------|------|------|
| Common sense | 1 | .04 | .96 |
| | 2 | -.11 | 1.04 |
| | 3 | .04 | 1.07 |
| Own experience | 1 | -.44 | .66 |
| | 2 | -.11 | .88 |
| | 3 | .62 | 1.16 |
| Theory | 1 | -.26 | .48 |
| | 2 | .69 | 1.36 |
| | 3 | -.02 | .92 |
| Error | 1 | .47 | 1.23 |
| | 2 | -.23 | .69 |
| | 3 | -.37 | .67 |

Further investigation based on Multiple Analysis of Variance (MANOVA) reveals how the mental thought processes' effects differ between the three groups.

Common sense

Teaching level did not have a significant effect on using common sense as a mental thought process. The advanced students used mental thought processes to a somewhat less degree than the two other groups, but the MANOVA revealed that there were no significant difference between the three groups ($F_{(2,42)} = .10, p = .902$).

Own Experience

Experienced teachers used a significantly higher proportion of mental thought processes based on Own Experience ($F_{(2,42)} = 4.67, p = .007$) than the two other groups. The Standard Deviation, however, indicates that there's a considerable variability among the experienced teachers ($SD = 1.16$). Advanced students used the category to some extent more than the beginning students.

Theory

Advanced students had a significantly higher proportion of mental thought processes based on Theory compared to the two other groups ($F_{(2,42)} = 3.33, p = .028$). However the Standard Deviation ($SD = 1.35$) implies the largest variability among all groups within each factor which points towards a substantial variability. Relying on reasoning based on Theory was followed by experienced teachers and was less frequently used among beginning students.

Error

The beginning students used a significantly higher proportion of mental thought processes based on Error compared to the two other groups ($F_{(2,42)} = 3.33, p = .032$). As well in this factor the Standard Deviation indicated a considerable variability (1.23). Reasoning based on Error was followed by experienced teachers, advanced students were the group which relied least on mental thought processes based on Error.

As stated in the purpose of the study we also wanted to investigate whether any of the factors predicted the overall performance on the test. Multiple regression analyses were performed with the overall test scores as a dependent variable. The results indicated that the four categories of mental thought processes explained a significant amount of the variance in the scores on the items from the Arizona Educator Proficiency Assessment (AEPA), $R^2 = .31, F(4, 40) = 4.51, p = .004$. A significant positive relationship was found for Common Sense, $\beta = .31, p = .028$. These results indicated that the test is not measuring pedagogical

knowledge and skills, but instead measures common sense. And as stated earlier, there was no significant difference between the groups when it came to relying on common sense when responding to the items.

The only significant factor which negatively predicted the AEPA test score was reasoning based on Error, $\beta = -.46$, $p < .001$. The participants which used mental thought processes which were categorised into this factor seemed to hesitate a lot more than participants which seemed to rely on other processes. A lot of the coded segments were simply repeating the question or alternative rehearsing until they had to choose an answer.

Table 6: Summary of regression analysis for variables predicting AEPA test score

| Variable | <i>B</i> | <i>SE B</i> | <i>B</i> |
|----------------|----------|-------------|----------|
| Common sense | .42 | .18 | .31* |
| Own experience | -.23 | .18 | -.17 |
| Theory | .02 | .19 | .01 |
| Error | -.63 | .18 | -.46*** |

* $p < .05$. ** $p < .01$ *** $p < .001$

A typical series of segments coded into this category would be:

“.... well hmm, let me read the question...”

“...Literature should not be “used” in the classroom; rather it should be “received by children. That is, literature is not simply a resource, a thing to practice on, but is of itself an experience to be entered into, to be shared and contemplated. This is what we must teach children, by discovery, that literature is...”

“...I just don’t know....”

“...let me read the alternatives again...”

“...A. It is an argument in favour of skill based reading programs...”

“...B. It is an argument in favour of allowing teachers to select the reading series used in their classrooms...”

“...C. It is an argument embodying many of the principles of whole-language instructional approaches...”

“...D. It is an argument against meta-cognitive approach to reading instruction...”

“...this is a difficult one...”

“...maybe C is the answer, it just seems right...”

“... or maybe B could be the answer...”

“... D, I don't even know what meta-cognitive approach to reading is....”

“...I guess I'll just have to pick one...”

“...I'm gonna go ahead and say alternative B for this one....”

The examples above are taken from the protocol in our study and are representative examples of mental thought processes relying on Error. The beginning students were the group which had the highest proportion of segmented processes based on error and the experienced teachers had the lowest proportion of reasoning based on Error.

Theory is the only factor that is demonstrating pedagogical knowledge and skills, where a typical segment would look like this:

“... according to Behaviourism this is a positive reinforcement which has a positive effect on children's learning...”

“... this is also consistent to latest research which are arguing pro the same statement...”

The advanced students did have a significantly higher proportion of reasoning based on Theory. Advanced students are these tests' target group, which can help validate the test.

7. Discussion

Our attempt was to understand the types of mental processes used in taking these kinds of tests. The study offer insight into the mental processes at work in individuals as they respond to items such as the AEPA practise test items used in our study. The method of verbal protocol data appears to be well suitable to this type of task and can provide an inside to the mental thought processes at use while responding to items similar to the ones we used. Most of the participants seemed to be comfortable thinking aloud while answering to the selected items.

The study demonstrates that it is possible to create a classification system for the segmented mental thought processes used in this type of studies as provided by Chi (1997). Participants in the study were engaged in a variety of mental processes, from using mental thought processes based on pedagogical knowledge and skills drawn from learning theories and research to using thinking not clearly productive of problem solutions, including total mental wandering, not at all relevant for the test.

The overall performance on the test indicates the selected items used in this study were of a moderate complexity, presenting frequent cues to call up associations from long-term memory into short-time memory where they are available for verbalizations. As Branch (1994) points out, problem can occur if the task involve a high cognitive load and if the process is automatic for the participant.

- *Which mental thought processes are most frequently used while responding to the selected items?*

The result demonstrates how frequently the mental thought process deciding about option were used, almost a quarter of all the segmented thought processes could be categorised into this category. This category was followed by repeating the question, rehearsal cues and common sense. These four mental thought processes can count for almost seventy percent of the total segmented protocols. The relatively large proportions of these categories

demonstrate that there is not a lot considered pedagogical thinking. This is consistent with Oberg's (2005) result who claims considered decision making processes only count for a small part of teachers' practise. He further says teaching is very complex, therefore teachers often act routinely or habitually. Hence, our results might point towards that participants made inferences from beliefs or intuition rather than evidence. Shavelson et al. (1977) also assumed that subjects' beliefs about education and teaching is expected to influence their judgments, decisions and behavior. Given that teachers has to make decisions on the spot, adds complexity to the situation. Acting routinely or by beliefs might be necessary.

- *Are any of the mental processes related to performance on the test?*

High performance on the selected items was associated with a greater proportion of construct relevant thinking process. Theory and Common Sense were the two predictors which could predict a score in the positive direction; however Common Sense was the only significant predictor. Buchmann's (1987) research on teaching knowledge argues the knowledge we use while teaching is not much different interacting with others on a daily bases. This means the common sense structures we use everyday obtain a big position in teachers' thinking. A teacher's profession is complex and you have to make difficult decisions on a constant basis. A thoughtful teacher making decisions might need a high proportion of common sense to handle all these situations. Still the items used in our study are supposed to measure pedagogical knowledge (AEPA). Nevertheless the results obtained by us can indicate that it is measuring common sense instead.

Further earlier research found that high achievement test scores is characterized by rapid judgment, reducing many events and cues into a small amount of categories and willingness to change the course when needed. Also much of teachers thinking are about the pupils' and the pupils' needs. However we used a multiple choice test was the participants were instructed to think aloud only concerning each alternative. Using additional methods like observing and interview might have given us further information. Still we used think aloud, hence it is difficult to find their willingness to change course when needed. Since the participants were instructed to think aloud to every alternative it is also hard to know the amount of thoughts actually concerning their pupils in the classroom.

- *Are there any differences between beginning students, advanced students and experienced teachers' mental thought processes while responding to the selected items?*

We did find differences in mental thought processes among the three groups. Experienced teachers did rely more on own experience. Advanced higher proportion of mental thought process relying on theory, but the overall proportion was extremely low. Beginning students used more mental processes based on error than the two other groups. Error was the only factor which had a significant negative prediction on test score. Low performance on the selected items was associated with non-resolution productive thinking, similar to the mental thought process categories classified into the factor Error. Using construct irrelevant strategies was associated with low performance on the test. According to Glaser (2005) novices might have the knowledge, but difficulties knowing how to apply this knowledge. This might reflect the mental thought processes used by the beginning students in our study. On the other hand you have expert or experienced teachers. According to research (e.g. Leinhardt, 2005 and Leinhardt and Greeno, 1986), they are more flexible while teaching. One believes this is because they have much better representation of the problem and can base their solution from such representations. The experienced teachers in our study had a high proportion of mental processes based on own experience. One can assume better representation of a problem comes from more experience.

What do these tests measure?

Another issue is licensure tests validity. Licensure tests are intended to measure pedagogical knowledge and skills, but if we look at the four most frequent thought processes none of these require any pedagogical knowledge. The results obtained in our study indicate that these tests are not measuring pedagogical knowledge and skills, but instead common sense.

However the advanced students obtained a significantly higher proportion of mental thought processes relying on these skills, which can help validate the test. Licensure tests, similar to the items we used in our study, are supposed to determine competent from incompetent

applicant before they start their teaching carrier. Consequently these tests are designed to measure students in their final period of study as they are about to start their professional carriers as teachers. In that context the test can be regarded as significant, advanced students did have a significantly higher proportion of mental thought processes relying on learning theories and research. Still the overall proportion was extremely low, and the standard deviation indicates that there was a considerable variation among the participants.

7.1 Limitation

The study had some limitations; the most obvious one has to do with validity, because of its sample size and sampling procedure. There were forty-five participants in the study, and when we used teacher level as a factor, the groups consisted of twelve, fifteen and eighteen participants. Because of the number of participants in each group, the results achieved in our study may be highly influenced by sampling error. On the other hand the laborious nature of these kinds of study, a large enough sample size to obtain significant estimates of infrequently used strategies might be difficult to achieve. With a larger sample size additional factors could as well reach statistical significance.

The participants in our study were also paid twenty dollars to participate in the study, which can influence the sample. It's difficult to know if participants who get paid to be in a study have different qualities than if the sampling would have been selected randomised.

Another major limitation in our study has to do with its reliability. The only consistency check was conducted by the same person; the transcript was not coded by a second-rater to assess inter-rater agreement on coding, which obviously affects its reliability.

There was no time limit in our study, however a few of the participants still rushed in some of the answers, an example from the study is given below:

“....I’m gonna try to hurry up...”

“....I know it’s late, and you probably want to go home soon...”

This segment might illustrate that even if there was no time limit, some of the participants did respond as if it was one. To have breaks when the participants were tired could have been an implication to improve the study.

8. Conclusion

Our attempt was to better understand teachers' mental processes. To discover this we investigated 45 beginning students, advanced students and experienced teachers. The results obtained in our study demonstrate that there is not a lot of considered pedagogical thinking. This is consistent with Oberg's (2005) findings, who argue considered decision making processes only count for a small part of teachers' practise. Common Sense was the only significant predictor which could predict a score in the positive direction. We also found a difference between the three groups in our study. Experienced teachers tended to rely more on own experience. Advanced higher proportion of mental thought processes relying on theory, but the overall proportion was very low. Beginning students used more mental processes based on error, which is consistent with some novice- expert studies.

The study do have some limitations, however the results obtained on our study can help us to understand the mental thought processes in use while answering these kinds of tests. The study revealed that there was not much pedagogical thinking, though advanced students did have a significantly higher proportion of mental thought processes relying on Theory which can provide evidence in support of the validity of the licensure test. Our results do however indicate that these type of tests measure common sense instead of pedagogical knowledge and skills.

Analysis of verbal data appears to be helpful to estimate relative effect, but perhaps not as an absolute estimation.

9. References

- Aanstoos, C.M. (1983). The think aloud method in descriptive research. *Journal of Phenomenological Psychology*, 243-266
- Arizona Educator Proficiency Assessment (AEPA). Retrieved from worldwide web on April 29, 2007, at www.aepa.nesinc.com
- Barr, R. (1975). Instructional Pace Differences and Their Effect on Reading Acquisition, *Reading Research Quarterly*, 9, 526-554.
- Berliner D.C. (2005). The Place of Process-Product Research in Developing the Agenda for Research on Teacher thinking. In P. M. Denicolo and M. Kompf (Eds.), *Teacher Thinking and professional Action*. (pp. 3-16), New York: Routledge.
- Berliner D.C. and Carter, K.J. (1986). *Differences in processing classroom information by expert and novice teachers*. Paper presented at the meeting of international Study Association on Teacher Thinking (ISATT), Leuven, Belgium.
- Bloom, B.S., & Broder, L.J. (1950). *Problem-solving processes of college students*. Chicago: University of Chicago Press.
- Branch, J.L. (2000) The Trouble with Think Alouds: Generating Data Using Concurrent Verbal Protocols. CAIS 2000: *Dimensions of a Global Information Science*. Canadian Association for Information Science Proceedings of the 28th Annual Conference.
- Buchmann, M. (1987) *Teaching Knowledge: The lights teachers live by*. Oxford Review of Education 13(2):151-64
- Calderhead, J. (1986). Developing a framework for the elicitation and analysis of teachers' verbal reports. Paper presented at the annual meeting of American Educational Research Association. San Francisco.
- Carnahan, R.S. (1980). *The effects of teacher planning on classroom processes* (Technical Report No. 541). Madison, WI: University of Wisconsin, Wisconsin Research and Development Center.
- Champagne, A.B., Gunston, R.F. and Klopfer, L.E. (1983). *A perspective on the differences between expert and novice performance in solving physics problems*. Pittsburgh: University of Pittsburgh, Learning Research and Development Center.
- Chase, W.G. and Simon, H.A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55-81.
- Chi, M.T.H. (1997). Quantifying qualitative analysis of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6, 271-315
- Chi, M. T. H., Feltovich, P.J. and Glaser, R. (1981). Categorization and representation of physics problems by expert and novices. *Cognitive Science*, 5, 121-152.

-
- Chung, G.K.W. K., de Vries, L. F., Cheack, A. M., Stevens, R. H., & Bewley, W. L. (2002). Cognitive process validation of an online problem solving assessment. *Computers in Human Behaviour*, 669-684.
- Clandinin, D.J. and Connelly, F. M. (1985). Rhythms in teaching: The narrative study of teachers' personal knowledge of classrooms. *Teaching and Teacher Education*, 2(4), 377-387.
- Clark, C., & Lampert M. (1986). The study of teacher thinking: Implications for teacher education. *Journal of Teacher Education* , 37 (5), 27-31.
- Clark, C. M., and Peterson, P. L. (1986). Teachers' thought processes. In Wittrock, M. C. (ed.), *Handbook of Research on Teaching*, Macmillan, New York, pp. 255–296.
- Clark, C.M. and Yinger, R.J. (1979). *Three studies of teacher planning* (Research Series No. 55). East Lansing: Michigan State University.
- Colker, L. (1982). *Teachers' interactive thoughts about pupil cognition*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.
- Connors, R. D. (1979). *An analysis of teacher thought processes, beliefs, and principles during instruction*. Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational measurement*. Washington, DC: American Council on Education.
- Dershimer G.M. & Oliver B. (2005). Examining complexity of thought in secondary student teachers. In P. M. Denicolo and M. Kompf (Eds.). *Teacher Thinking and professional Action* (pp. 211-222), New York: Routledge.
- Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37, 15-27.
- Elbaz, F. (1983) *Teacher thinking: A study of practical knowledge*. London: Croom Helm.
- Erickson, F. (1984). What difference does teaching experience make? *IRT Communication Quarterly*,
- Ericsson, K. A. (1987). Theoretical implication from protocol analysis testing and measurement. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology on testing*. Hillsdale, NJ: Erlbaum.
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev.ed.). Cambridge, MA: MIT Press.
- Hamilton, L. S., Nussbaum, E. M., & Snow, R. E. (1997). Interview Procedures for Validating Science Assessments. *Applied Measurement in Education*, 181-201.
- Hidi, S. and Klainman, R. (1983). Note taking by experts and novices: an attempt to identify teachable strategies, *Curriculum Inquiry*, 13, 397-418.

Hill, J., Yinger, R. J. and Robbins, D. (1981). *Instructional planning in a development preschool*. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.

Huber G.L. (2005). Strategies and methods on teacher thinking. Routledge, New York In P. M. Denicolo and M. Kompf (Eds.). *Teacher Thinking and professional Action* (pp. 61-76), New York: Routledge.

ISSAT, from worldwide web on April 29, 2009, at www.issat.org.

Kompf M. & Brown A.F. (2005). Teachers' personal and professional ideals about practise. In P. M. Denicolo and M. Kompf (Eds.). *Teacher Thinking and professional Action* (pp. 99-108), New York: Routledge.

Leighton, J. P. (2005). Avoiding misconception, misuse, and missed opportunities: The collection of verbal reports in educational achievement testing. *Educational measurement: Issues and practice*, 6-15.

Leinhardt, G. (2005). A contrast of novice and expert competence in math lessons. In P. M. Denicolo and M. Kompf (Eds.). *Teacher Thinking and professional Action* (pp. 41-58), New York: Routledge.

Leinhardt, G., and Greeno, J. G. (1986) The cognitive skill of teaching. *Journal of Educational Psychology*, 78(2),75-95.

Lissmann U. (2005). Analysing teachers' thoughts prior to student assessment. In P. M. Denicolo and M. Kompf (Eds.). *Teacher Thinking and professional Action* (pp. 139-152), New York: Routledge.

Marland, P. W. (1977). *A study of teachers' interactive thoughts*. Unpublished doctoral dissertation, University of Alberta, edmonton, Canada.

Marx, R. W., and Peterson, P. L. (1981). The nature of teacher decision making. In B. R. Joyce, C. C. Brown, and L. Peck (Eds.), *Flexibility in teaching: An excursion into the nature of teaching and training*. New York: Longman.

McNair, K. (1978-1979). Capturing inflight decisions. *Educational Research Quarterly*, 3(4), 26-42.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement*. Washington DC: American Council on Education & National Council on Measurement in Education.

National Institute of Education, 1975. National Institute of Education: Career Education Program: Program Plan for FY 1975.

National Research council (U.S). Committee on Assessment and Teacher Quality. Karen J.Mitchell (et al.), editors. (2001) *Testing teacher candidates: the role of licensure tests in improving teacher quality / Committee on Assessment and Teacher Quality*, Centre for Education, Board on Testing and Assessment, Division on Behavioural and Social Sciences and Education, National Research Con.

-
- Newell, A., and Simon H. A. (1956). The logic Theory machine. *IRE Transactions on Information Theory*, 1956, 2, 61-69.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 231-259.
- Oberg A. (2005). The ground of professional practise. In P. M. Denicolo and M. Kompf (Eds.). *Teacher Thinking and professional Action* (pp. 77-90), New York: Routledge.
- Olson J. (2005). Case study in research on teaching: A ground for reflective practise. In P. M. Denicolo and M. Kompf (Eds.). *Teacher Thinking and professional Action* (pp. 91-98), New York: Routledge.
- Payne, J. W. (1994) Thinking Aloud: Insights Into Information Processing. *Psychological Science*, vol 5. no 5, 241-248.
- Rutter, M., Maughan, B. Mortimor, P., Ousten, J. And Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge, MA: Harvard University Press.
- Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory and Cognition*, 759-769.
- Simon, Herbert A., (1979). Rational decision making in business organizations. *American Economic Review* 69, pp. 493-513.
- Shavelson, R. J. (1973). *The basic teaching skill: Decision making* (R & D Memorandum No. 104). Standford, CA: Standford University, School of Education, Center for R & D in Teaching.
- Shavelson, R. J. (1976). Teachers' decision making. In N. L. Gage (Ed.), *The psychology of teaching methods* (Yearbook of the National Society of the Study of Education). Chicago: University of Chicago Press.
- Shavelson R.J., (1981). Research on Teachers' Pedagogical Thoughts, Judgments, Decisions, and Behavior. *Review of Educational Research Vol. 51, No. 4, Pp. 455-498*. University of California, Los Angeles and the Rand Corporation.
- Shavelson, R. J., Atwood, N. K., and Borko, H. (1977). Experiments on some factors contributing to teachers' pedagogical decisions. *Cambridge Journal of Education*, 7, 51-70.
- Shavelson, R. J. and Stern, P. (1981). Research on teachers' pedagogical thoughts, judgements, decisions, and behaviour. *Review of Educational Research*, 51, 455-498.
- Shulman, L. S. and Elstein, A. S. (1975). Studies of problem solving, judgment, and decision making: Implications for educational research. In F. N. Kerlinger (Ed.), *Review of Educational Research*, 51, 455-498.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.

Smith I.L., and Hambelton, R (1990). Content validity studies of licensing examinations. *Educational Measurement: Issues and practise*, v9 n4 p7-10 Win 1990.

Taylor, P. H. (1970). *How teachers plan their courses*. Slough, Bucks, England: National Foundation for Educational Research.

Wade S. E. (1990). Using think alouds to assess comprehension. *The Reading Teacher*, vol 43. no. 7, 442-451.

Williams G. & Clarke D. (2002). The contribution of student voice in classroom research. *South African Association for Research in Mathematics, Science and Technology Education Conference*, University of Natal, Durban, South Africa, 398-404.

Yinger, R. J. and Clark, C. M. (1982) *Understanding Teachers' Judgements about Instruction: the task, the method, and meaning*, (East Lansing, MI., Institute for Research on Teaching).

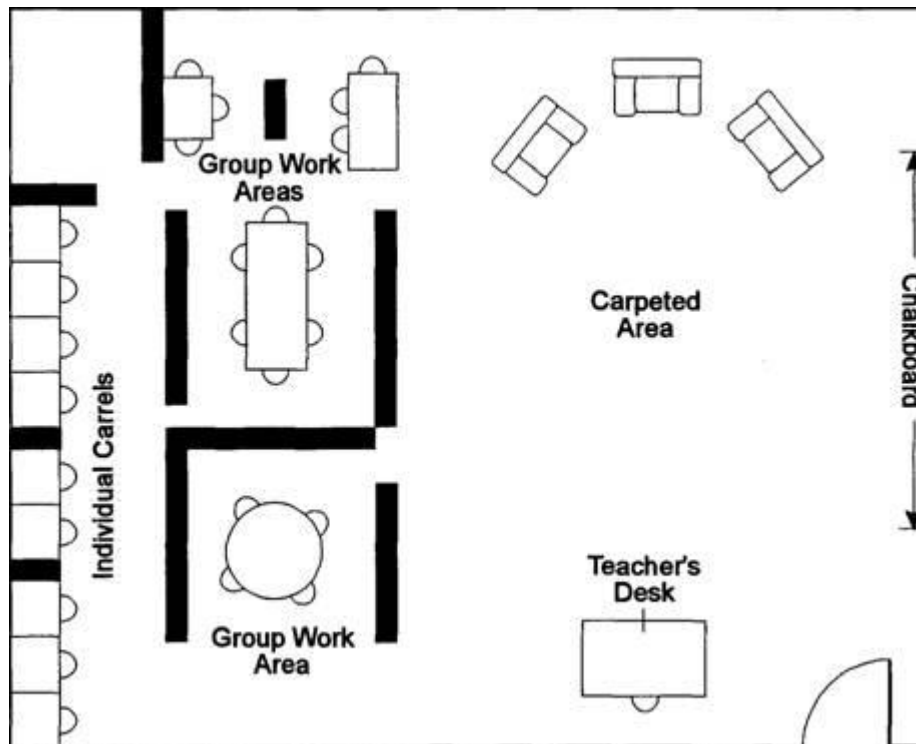
Zahorik, J. A. (1975). Teachers' planning models. *Educational Leadership*, 33, 134-139.

10. Appendix

Form A

EXAMPLE

Use the floor plan below of an elementary school classroom to answer the question that follows.



This classroom arrangement is most likely designed primarily to:

- A. accommodate the preferences and needs of students with a variety of learning styles.
- B. communicate the idea that learning in group contexts has advantages over learning independently.
- C. facilitate instruction that integrates concepts and materials from different disciplines.
- D. enable the teacher to monitor off task student behavior easily and effectively.

Directions: Each of the multiple-choice questions below is followed by four suggested answers or completions. Select the one that is the **best** answer.

1. Classroom management research findings suggest that one of the most effective ways to maximize the amount of time elementary school children spend on academic activities is for the teacher to do which of the following?

- A. Plan for, teach, and enforce routines for transition times and classroom housekeeping tasks
- B. Assign homework three times a week in the major subjects
- C. Assign individual reading on new topics before discussing the topic in class
- D. Introduce new material in a lecture followed immediately by a questioning session on the material

2. Literature should not be "used" in the classroom; rather it should be "received" by children. That is, literature is not simply a resource, a thing to practice on, but is of itself an experience to be entered into, to be shared and contemplated. This is what we must teach children, by discovery, that literature is.

Which of the following best characterizes the view presented in the passage?

- A. It is an argument in favor of skill-based reading programs
- B. It is an argument in favor of allowing teachers to select the reading series used in their classrooms
- C. It is an argument embodying many of the principles of whole-language instructional approaches
- D. It is an argument against a meta-cognitive approach to reading instruction

3. During a visit to a second-grade classroom, a student teacher observed a child spending the time allotted for a worksheet either looking out the window or doodling on his paper. When the student teacher asked the child if he needed help on the assignment, he said no. When asked why he wasn't doing it, he pointed to another student and said, "She does all her work fast and when she's done, she gets more work."

The boy's reaction suggests which of the following about his classroom?

- A. A routine has been established for students who are having trouble finishing an assignment to ask the teacher for assistance.
- B. A routine for rewarding students who finish work promptly is not in place.
- C. Students must work alone on seatwork, without consulting other students.
- D. Students who finish work before the whole class is finished must not interrupt the students who are still working

Please continue to the next page.



4. Decide whether the action makes it likely or unlikely that the goal will be achieved and select the best statement of the reason that the action is likely or unlikely to lead to the achievement of the goal.

GOAL: To encourage a student to perform at a higher level than he had been

ACTION: Give the student a somewhat higher grade than the quality of the student's work warrants

- A. LIKELY, because the student will feel more self-pride than if the grade had been commensurate with the work
 - B. LIKELY, because the student will conclude that better performance will elicit an even higher grade
 - C. UNLIKELY, because the student will realize that the reward is out of proportion to the accomplishment
 - D. UNLIKELY, because the student may feel that the current effort is sufficient
-

5. Which of the following activities would require students in a social studies class to apply the most advanced level of cognitive skills?

- A. Students read a passage about the multiple causes of an historical event and then discuss how each cause contributed to the event.
- B. Students listen to a presentation about a series of historical events and then create a timeline based upon that information.
- C. Students select an historical event and then use a print or non-print resource to find information about the event.
- D. Students identify possible explanations for an historical event and then discuss types of evidence that would support each explanation.

6. A teacher whose students achieve at a range of levels is considering using ability grouping for much of the instruction in his classroom. According to educational research, such an approach is most

likely to have which of the following **negative** effects?

- I. increasing obstacles to achieving a sense of community and mutual support among all students
 - II. reducing all students' sense of achievement in regard to their own learning
 - III. communicating low expectations for students who currently perform at lower levels of achievement
 - IV. making it more difficult to individualize instruction for students with varied needs
- A. I and III only
 - B. I and IV only
 - C. II and III only
 - D. II and IV only



7. A teacher begins an ecology unit by asking her students to discuss what the term *ecology* means to them. Which of the following are the primary benefits of introducing the unit in this way?

- I. The teacher will get a better idea of how to group students during activities related to unit.
- II. Students will have a chance to reflect on what they already know about the topic.
- III. The teacher will obtain information about students' current understanding of the topic.
- IV. Students will be encouraged to work collaboratively during the remainder of the unit.
 - A. I and III only
 - B. II and III only
 - C. I and IV only
 - D. II and IV only

8. A teacher records on the table below the grades earned by his students on assignments and quizzes addressing different topics in a unit.

| | Average Assignment Score (%) | Average Quiz Score (%) |
|----------------|---|-----------------------------------|
| | 74 | 72 |
| Topic 2 | 82 | 86 |
| | 61 | 56 |
| Topic 4 | 90 | 87 |

Which of the following is the most appropriate conclusion to draw from this information?

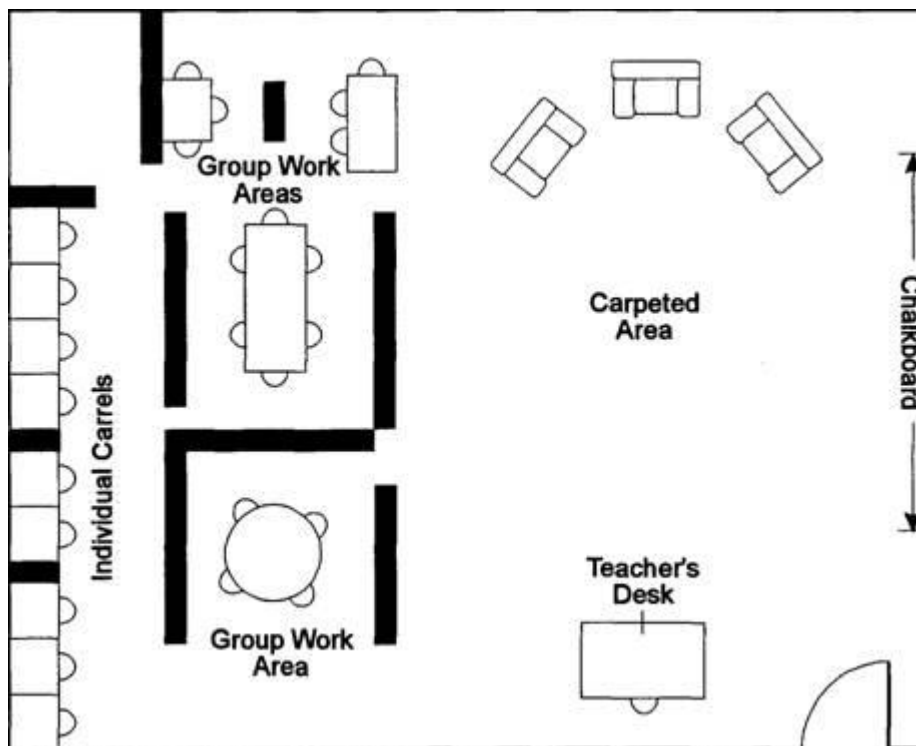
A. Assignments were too easy and failed to prepare students adequately for quizzes. Additional instruction should be provided to reinforce Topic 3 before moving on to a new unit.

- B. Substantial variation in the quality of instruction during the unit led to inconsistent student performance.
- C. Instruction related to Topic 4 did not provide sufficient challenge for the students in this class.

-- Stop. You have completed PHASE ONE. --

Form B**EXAMPLE**

Use the floor plan below of an elementary school classroom to answer the question that follows.



This classroom arrangement is most likely designed primarily to:

- A. accommodate the preferences and needs of students with a variety of learning styles.
- B. communicate the idea that learning in group contexts has advantages over learning independently.
- C. facilitate instruction that integrates concepts and materials from different disciplines.
- D. enable the teacher to monitor off task student behavior easily and effectively.

1. A math teacher often begins class periods by selecting students to go to the chalkboard to write out and explain their solutions to homework problems. If a student cannot solve a problem or has trouble explaining a solution, the teacher requests that a classmate volunteer

to assist. The most important **disadvantage** of this instructional approach is that it is likely to:


- A. communicate to all students that the teacher has low expectations for their academic performance and progress.
- B. focus too much student attention on learning processes and too little attention on learning products.
- C. have an adverse effect on some students' self-esteem and on the overall emotional climate of the classroom.
- D. undermine students' willingness and ability to use more cooperative approaches in other class activities.

2. At the beginning of the year, a fifth grade teacher has a "Get to Know You" conference with each of his students. One of the students has a physical condition that restricts her ability to hold and manipulate books and other materials. Which of the following is the most appropriate way for the teacher to communicate sensitivity to this student's special needs?

- A. Assure her that although she will be given the same assignments as her peers, allowances will be made for her disability in grading some aspects of her work.
- B. Avoid raising the topic of her disability and downplay its significance if she expresses concern about it.
- C. Acknowledge her disability and offer to work with her to adapt class activities to make sure she has every opportunity for success.
- D. Tell her about other students with physical disabilities who have had successful experiences in the teacher's classroom.

3. A kindergarten class recently visited a local post office, where a postal worker showed the children how the post office functions and talked to them about his job. Which of the following steps taken by the teacher would best build on the post office experience to promote children's continuing exploration of the world of work?

- A. encouraging students to notice the many other types of workers they regularly encounter in their community
- B. setting up in the classroom a "post office" that is stocked with envelopes, stamps, and mail bags
- C. holding a class discussion about the various types of work done by their own parents/guardians
- D. having each student dictate to the teacher an account of the post *office* trip for display in the classroom

Please continue to the next page. 

4. A teacher plans to assign his students a research project and asks the school librarian to provide the class with instruction on library use. The teacher can best help ensure that the library session will be productive by providing the librarian with information about:

- I. students' current research and library skills.
 - II. the criteria the teacher will use to assess student work on the project.
 - III. motivational strategies that have been effective with this class.
 - IV. the nature of the research project the teacher will be assigning.
- A. I and III only
 - B. I and IV only
 - C. II and III only
 - D. II and IV only
-

Use the information below to answer the two questions that follow.

At the end of Bicycle Safety Week, Mr. Flint asks his first graders to draw pictures about what they have learned and to write about their pictures using invented spelling. Walking around the room 15 minutes later, he sees that Bethany has not made any attempt to write. The following dialogue takes place.

Mr. Flint: Remember to write a sentence about your picture, Bethany.

Bethany: I can't. I don't know how.

Mr. Flint: Sure you can. Just think of what you want to say. Then write down the sounds you can hear.

With some more coaxing from Mr. Flint, Bethany finally writes B SAF UN YR BIC below her drawing. Then she announces, "I know it's wrong. Dad says writing's hard for me, just like it was for him."

5. Based on this interaction, which of the following is the most reasonable conclusion for Mr. Flint to draw?

-
- A. Bethany should be evaluated to determine whether she has inherited a learning disability that affects her writing skills.
 - B. Bethany's father does not take a great deal of interest in the progress his daughter is making in school.
 - C. Bethany will need a more structured approach to writing instruction to develop competence as a writer.
 - D. Bethany's progress toward writing is being impaired because her father is communicating low expectations for her in that area.


Please continue to the next page.

6. To follow up on his concerns about Bethany, Mr. Flint invites her father to meet with him. As he begins a brief explanation of how he uses the whole language approach as part of his language arts instruction, Bethany's father becomes exasperated. "I just don't see enough evidence that Bethany is learning," he says. "I'd like to see more worksheets. I'd like to see more spelling tests." Which of the following responses would be most appropriate and effective for Mr. Flint to make?

- A. Present a portfolio or other examples of Bethany's work and point out specific ways in which she is making progress.
- B. Assure Bethany's father that much current research supports the effectiveness of the whole language approach.
- C. Explain that a child can learn spelling words and do skills practice with worksheets without becoming a good reader or writer.
- D. Offer to individualize Bethany's instruction by giving her more opportunities to complete skills worksheets.

7. Decide whether the action makes it likely or unlikely that the goal will be achieved and select the best statement of the reason that the action is likely or unlikely to lead to the achievement of the goal.

GOAL: To encourage a student to perform at a higher level than he had been

ACTION: Give the student a somewhat higher grade than the quality of the student's work warrants

- A. **LIKELY**, because the student will feel more self-pride than if the grade had been commensurate with the work
- B. **LIKELY**, because the student will conclude that better performance will elicit an even higher grade
- C. **UNLIKELY**, because the student will realize that the reward is out of proportion to the accomplishment
- D. **UNLIKELY**, because the student may feel that the current effort is sufficient

Directions: Each of the multiple-choice questions below is followed by the four suggested answers or completions. Select the one that is best in each case.

8. Classroom management research findings suggest that one of the most effective ways to maximize the amount of time elementary school children spend on academic activities is for the teacher to do which of the following?

- A. Plan for, teach, and enforce routines for transition times and classroom housekeeping tasks
- B. Assign homework three times a week in the major subjects
- C. Assign individual reading on new topics before discussing the topic in class
- D. Introduce new material in a lecture followed immediately by a questioning session on the material

Please continue to the next page.



9. During a visit to a second-grade classroom, a student teacher observed a child spending the time allotted for a worksheet either looking out the window or doodling on his paper. When the student teacher asked the child if he needed help on the assignment, he said no. When asked why he wasn't doing it, he pointed to another student and said, "She does all her work fast and when she's done, she gets more work."

The boy's reaction suggests which of the following about his classroom?

- A. A routine has been established for students who are having trouble finishing an assignment to ask the teacher for assistance.
- B. A routine for rewarding students who finish work promptly is not in place.
- C. Students must work alone on seatwork, without consulting other students.
- D. Students who finish work before the whole class is finished must not interrupt the students who are still working

-- Stop. You have completed PHASE ONE. --