

ASSESSING THE VALIDITY AND RELIABILITY OF A PIAGETIAN BASED PAPER-  
PENCIL TEST

A Thesis by

Amy Dugan

B.A., Wichita State University, 2003

Submitted to the Department of  
Counseling, Educational, & School Psychology and  
the faculty of the Graduate School of  
Wichita State University in partial fulfillment of  
the requirements for the Degree of  
Masters of Education

May 2006

ASSESSING THE VALIDITY AND RELIABILITY OF A PIAGETIAN BASED PAPER-  
PENCIL TEST

I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Education with a major in Educational Psychology.

---

Linda Bakken, Ph.D. Committee Chair

We have read this thesis and recommend its acceptance:

---

Marlene Schommer-Aikins, Ph.D. Committee Member

---

Sue Neal, Committee Member

## DEDICATION

To my parents, for without their love and support I could have never made it this far.

## ACKNOWLEDGMENTS

I would like to thank Dr. Bakken for serving as my thesis chair. Dr. Bakken went above and beyond the call of duty when helping me edit and revise my thesis. I would also like to thank Dr. Schommer-Aikins for her guidance when I encountered bumps and Sue Neal for serving as a member on my committee.

## ABSTRACT

Piaget hypothesized that we understand the world through the use of cognitive frameworks. The strengths and limitations of these cognitive frameworks (or levels) both help and hinder the learning process. Comprehending individuals' cognitive levels is essential for teachers to ensure that their students learn. Piaget developed a series of tasks to assess individuals' cognitive levels (typically called "clinical interviews" which can only be conducted in a one-to-one fashion). Thus, a paper-and-pencil test that could be administered to groups was developed to help teachers determine the cognitive level of the children they teach. Problems with the scoring technique limited the validity and reliability of the instrument; therefore, a revised scoring system was developed that simplified and broadened the scoring of the test. The purpose of the current study is to determine if the reliability and validity of the paper-and-pencil instrument would be significantly increased through the use of the revised scoring procedures. Pre-existing data will be used in all analysis. A total of 279 students (ranging from third to twelfth grades) took the paper-and-pencil test. Next, each student either completed the Piagetian tasks under the supervision of a trained task administrator in the traditional one-on-one format or was retested using the paper-pencil instrument. Some students participated in all three assessments. A bivariate correlation was conducted to analyze the validity and reliability of the instrument. A t-test was calculated to test for a significant difference in the correlational coefficients between the two scoring methods. Results show that correlation coefficients are stronger when using the revised method. The t-test found that the revised scoring method was significantly more reliable, yet was only more valid for two of three stages.

## TABLE OF CONTENTS

Chapter	Page
1. THE PROBLEM	1
Purpose	2
Definitions	3
Overview	4
2. LITERATURE REVIEW	5
Piaget's Theory	5
Sensorimotor Stage	7
Preoperational Stage	8
Concrete Operational Stage	9
Formal Operational Stage	11
Research of Theory	15
Testing Children for Cognitive Development	18
Conservation of Mass	19
Combinatorial System of Operations	20
Paper-and-Pencil Tests	20
Summary	24
Hypotheses	24
3. THE METHOD	26
Participants	26
Instrument	26
Procedure	30
Data Analysis	31
4. THE RESULTS	32
Reliability	32
Validity	37
5. DISCUSSION	42
Limitations	43
Implications for Future Research	43
Summary	44

REFERENCES	45
APPENDICES	
Appendix A: Description of Piagetian Tasks	49
Appendix B: Bakken's Paper and Pencil Test	51

## LIST OF TABLES

Table	Page
1. Summary of Validity Coefficients Reported by Skar, et. al., 1995	28
2. Summary of Reliability Coefficients Reported by Skar, et. al., 1995	29
3. Original and Revised Scoring Methods for Piaget Paper-Pencil Test	31
4. Means and Standard Deviations for Reliability Coefficients for Third, Fifth and High School Students on a Piagetian Paper-and-Pencil Test Using Original and Revised Scoring	33
5. Reliability Coefficients for Third, Fifth, and High School Students on a Piagetian Paper-and-Pencil Testing Using Original Scoring and Revised Scoring	34
6. Reliability Coefficients between Original and Revised Scoring Methods for Third, Fifth, and High School Students on a Piagetian Paper-and-Pencil Test Using Original scoring and Revised Scoring	36
7. Means and Standard Deviations for Validity Coefficients for Third, Fifth, Seventh, and High School Students on a Piagetian Paper-and-Pencil Test and a Piagetian based Clinical Assessment Using Original scoring and Revised Scoring	37
8. Validity Coefficients for a Piagetian Paper-and-Pencil Test and Piagetian based Clinical Assessment Scores for Third, Fifth, Seventh and High School Students using Original Scoring and Revised Scoring	39
9. Validity Coefficients between original and revised scoring methods for Third, Fifth, Seventh, and High School Students on a Piagetian Paper-and-Pencil Test using Original and Revised Scoring	41

## CHAPTER I

### THE PROBLEM

In 2004 37.2% of Wichita Public Schools tenth grade students scored proficient or above on the state math assessment; seventh grade math scores were not much greater: 48.2% (Funk, 2005). In other areas the rate of proficiency was similar with reading and writing scores for middle and high school students ranging from 53% to 65%. One reason that scores may be low for adolescent learners is the disconnect between students' cognitive development and the expectations for curriculum by school districts. The way students learn and processes information is dependent upon their cognitive ability (Crain, 2005). It is imperative to know a person's cognitive ability so that material can be presented to the students in ways that are appropriate for the individuals' levels of cognition; people cannot learn information that is presented in a manner that is too complex for their level of cognitive understanding. If teachers knew in what stage of cognitive development their students' were operating, they could plan developmentally appropriate lessons. Theoretically, this would increase student learning. One means of determining students' learning capabilities would be to assess their level of cognitive development.

Piaget hypothesized that children learn in qualitatively different ways based upon their level of development. He placed people in one of four stages: sensorimotor, preoperational, concrete, and formal operations (Muuss, 1996). In order to determine in what stage of cognitive development a person was operating, Piaget developed a series of tasks which he used to assess children's levels of cognitive abilities. The clinical method developed by Piaget is administered in a one-on-one setting; it is not a method which is

feasible for use by a classroom teacher. Thus, there is a need for a paper-pencil test which could be administered by classroom teachers to groups of students.

### *Purpose*

The development of a Piagetian-based paper pencil test will allow educators to assess the cognitive abilities of their students. The No Child Left Behind Act of 2002 increased the accountability of educators. In order for students to be more successful in understanding and retaining what their teachers are teaching, the material must be presented to them in ways that are cognitively appropriate. The traditional method (clinical interviewing) of assessing a students' cognitive level is not feasible to the classroom teacher due to time constraints and lack of training. A paper-pencil test which correctly identifies the cognitive abilities of students would allow classroom teachers to assess the cognitive level of each student in one sitting. This would enable teachers to focus their time and energy developing lessons that are theoretically based.

Bakken (1995) created a paper-pencil test based on Piaget's tasks. The test can be used by classroom teachers who wish to determine the stage of cognitive development each of their students. Unfortunately, problems with the scoring technique limited the validity and reliability of the instrument; therefore, a revised scoring system was developed that simplified and broadened the scoring of the test. The purpose of the current study was to determine if the reliability and validity of the paper-and-pencil instrument would be statistically significantly increased through the use of the revised scoring procedures.

The researcher used existing data to analyze whether or not the revised scoring method statistically, significantly increased the reliability and validity of the instrument.

The data the researcher used was collected by various researchers in the late 1990s. Researchers collected the data at three schools in a metropolitan area. One of the schools was a dual-language magnet; it has a high percentage of Hispanic, ESL students, as well has a high percentage of students who come from a low SES household. The other two schools were parochial schools.

### *Definitions*

As children learn and process information, they are continually going through a process where they are restructuring new information and restructuring their minds to store the new information (Muuss, 1996).

*Assimilation* is the process which information is changed to fit into the child's existing mental structure (Crain, 2005).

*Accommodation* is the process in which changes are made in the mental structure so that new information can be stored and utilized. Accommodation and assimilation are complementary processes that interact to bring about conceptual adaptation and cognitive growth (Muss, 1996, p.147).

As children make accommodations and assimilate new information, there is harmony between the sensory information and accumulated knowledge; during this time children are in a stage of *equilibrium*. According to Piaget (1967) equilibration is "an overriding principle of mental development in the sense that all mental growth progresses toward ever more complex and stable levels of organization" (p.1). When children are in a state of equilibration, their minds are comfortable because all information has been processed and accommodated or assimilated.

As children learn new information and concepts their mind enters into a state of *disequilibrium* (Muuss, 1996). It is in disequilibrium that they must accommodate and assimilate the new information in order to reach a new stage of equilibrium. It is the continual process that moves children to more complex levels of thinking. If a child never went through these processes, he would never develop more sophisticated ways of thinking.

#### *Overview*

The second chapter of this proposal provides a more detailed description of Piaget's theory of cognitive development. It also reviews the literature relevant to the development of paper-pencil testing of cognitive development based upon the tasks developed by Piaget. Chapter Three presents the design of the study including the sample and instrument; it will also summarize the method of data collection. Chapter Four will analyze the data in terms of the hypotheses. Chapter Five will discuss the meanings of the results sections, state the limitations of the study, and propose areas of future research.

## CHAPTER II

### LITERATURE REVIEW

In order to assess the need for a paper-pencil test based on Piagetian principles, one must have a basic understanding of Piaget's theory of cognitive development. Therefore, this chapter will begin with an overview of the theory. Although each stage of his theory will be discussed, the later stages of the theory, concrete operations and formal operations will be discussed in greatest detail because they are the most relevant to this study. Next, research regarding Piaget's theory of cognitive development will be introduced. Third, will be a discussion of Piaget's clinical interview method. Fourth, Literature that discusses paper-pencil tests will be reviewed. After a brief summary, the chapter will conclude by stating the researcher's hypothesis.

#### The Review

##### *Piaget's Theory*

Describing Jean Piaget's (1896-1980) theory of cognitive development as a maturational theory or an environmental theory would be an over simplification. A new branch of psychology, genetic epistemology was developed by Piaget (Elkind, 1967). In a 1967 conversation with Bringuier (1977) Piaget stated:

...if you study the formation of knowledge, which is my craft, you must constantly identify the intervening factors- those due to external experience, social life, or language and those due to the internal structure of the thinking of the subject, which is constructed as it develops... these are all related epistemological problems. (p.18)

Piaget surmised that these four factors intermingle to produce cognitive development within an individual (Crain, 2005). When questioned as to whether external or internal experiences were more important, Piaget responded by stating that the experiences were of equal importance and incapable of being separated from one another (Bringuier, 1977). Piaget strove to understand children's internal thought processes; this focus is different from the conventional study of cognition.

Developmental change occurs through the "interaction between external stimulation and the nature of the child's cognitive stage-structures" (Cowan, 1978, p.20-21). Basically, children become more knowledgeable when they are better able to organize new information. They build mental structures (schemas) in which they organize information when they encounter an external experience and attempt to organize the experience with information they already know. The development of these schemas helps children to progress from one stage to the next (Crain, 2005). It is the child's internal process of dealing with environmental factors that moves the child from stage to stage, not the environment itself.

The interaction that occurs between the environment and internal thought process was of such interest to Piaget that he included it in his theory. Piaget suggested that when a child encounters new experiences or new information that the child is not mentally prepared to deal with, disequilibrium occurs. In order to return to equilibrium, the child must mentally adapt himself/herself so that the new experiences and/or information can be processed (Muuss, 1988). This process occurs continuously throughout development and the restructuring of the mind takes the child from his/her current stage of development to a more complex stage.

One way to return to equilibrium is to transform the *information* to fit into existing mental structures. This process called assimilation, allows children to further develop knowledge they already have. If children had to make a new mental structure for every new piece of information, they would not be able to relate concepts to others or develop a more in depth understanding of a concept they already understood at a basic level (Cowan, 1978). In this process children are using prior knowledge to help them understand the new information.

Another structure children are able to use in order to enter into equilibrium after encountering new information is accommodation. Whereas during assimilation children transform *information*, during accommodation children transform their *mental structure*. In this process children are unable to assimilate the new information so they build more complex mental structures that the new and old information can fit into. When this has been accomplished, the child moves back into a state of equilibrium (Cowan, 1978).

Four periods of cognitive development were introduced by Piaget: sensorimotor, preoperational, concrete, and formal operations (Muuss, 1996). Each stage is based upon cognitive abilities and thought processes. The periods of development are continuous in that each stage builds upon the previous stage but discontinuous in that each stage is characterized by the development of cognitive structures that are qualitatively different than the preceding stage (Wadsworth, 1977). Movement between stages occurs in an invariant sequence and, although Piaget did not assert that children progress through stages at fixed times, he did hypothesize an age guideline which can be utilized when classifying children.

*Sensorimotor Stage.* The sensorimotor stage is the stage into which neonates are born. The sensorimotor stage has six substages: basic reflexes (birth to one month), primary circular reactions (one month to four months), secondary circular reactions (four months to eight months), coordination of secondary circular reactions (eight months to twelve months), tertiary circular reactions (twelve months to eighteen months), and invention of new means through mental combinations (18 months to 24 months) (Muuss, 1996). Children in this stage are governed by egocentrism and are developing hand-eye coordination (Muuss, 1996). The task of the child in the sensorimotor stage is to move from reflex activity to activity invoked by thought (Wadsworth, 1977).

*Preoperational Stage.* The second stage of development is the preoperational stage. Children in this stage are beginning to understand more complex issues but are still dependent on their senses and can only focus on one dimension at a time (Crain, 2005). The major limitation of this stage is that children's minds are unable to concentrate on more than one aspect of any phenomenon. They center on one aspect of anything they are presented with. Therefore, a major task of this stage is to develop the ability to focus on more than one aspect of a phenomenon.

The preoperational stage is broken into two substages; both cause children's minds to become disorganized thus forcing children to restructure the way they think. The first substage is labeled egocentric (2-4 years) and the second substage is labeled intuitive (5-7 years) (Wadsworth, 1977). The names of each substage are derived from the way the child's mind works at each stage. Children in the egocentric substage cannot understand another person's point of view (Crain, 2005). The intuitive stage is characteristic of thinking based on perception (Cowan, 1978). The central task of the

preoperational child is to develop rational solutions to concrete problems (Wadsworth, 1977).

*Concrete Operational Stage.* The third stage is the concrete operational stage; this is the stage that children begin to use logical thinking (Muuss, 1996). Children in concrete operations (ages 7-11) can deal with problems that are realistic in nature. They are able to fully understand what is occurring around them and they can understand possibilities based on realities, but they are unable to think abstractly (Boden, 1979).

Until this stage of development children are egocentric. The concrete operational child is sociocentric (Muuss, 1996). Children in this stage of development can understand another's point of view. This is often illustrated in conversation. Children in this stage of development will engage in conversations with others in an attempt to both enlighten and understand others. This is not to say that a person in concrete operations is never egocentric, but that their general mode of thought is not egocentric.

The concrete operational child has a better understanding of social hierarchy. Children in concrete operations make moral judgments differently than they did in earlier stages. Now they can understand reasoning behind a rule. The child now knows that rules are not absolute but are made to provide structure and can be changed or modified as needed (Crain, 2005). Due to this new realization, children in this stage may begin to question authority.

The concrete operational stage of development is characterized by four elements of logical thinking: (a) the logic of classes, (b) the logic of relations, (c) the principle of conservation, and (d) the reversibility of thought processes. At the same time "four concrete group-like structures" are beginning to emerge: (a) combinativity, (b)

reversibility, (c) associativity, (d) identity or nullifiability (Muuss, 1996, p.156). These group-like structures are the cognitive processes that would enable logical thinking. Neither the logical thinking nor the cognitive processes could occur without the other; they are developed simultaneously. Due to the development of logical thinking and the corresponding group-like structures children no longer think impulsively; yet they still do not think systematically.

Relationships between parts and wholes can now be understood. Children in this stage can tell if something belongs to a class; this enables children to classify (Muuss, 1996). For example, if a preoperational child was shown ten black-plastic buttons and two white plastic buttons and was then asked whether there were more black buttons or more plastic buttons, the child would say that there were more black buttons (he cannot think about color **and** material). In the same situation, a concrete operational child would be able to classify the buttons by color and material; therefore, this child would know that there were more plastic buttons.

Concrete operational children are able to use the logic of relations; this means that they can relate objects to one another given a concrete criterion. Concrete operations is the first stage in which children can demonstrate an understanding of relations. Due to the development of the logic of relations, children who are shown three objects can now describe where one object is in terms of another. For example, when looking at a shoe, a straw, and a hat (in this order from left to right), the child could report that the straw is to the right of the shoe and to the left of the hat. He could also reverse the relationships; the shoe is to the left of the straw, and the hat is to the right of the straw.

The principle of conservation begins to be understood. Children are able to first conserve mass, then weight. Some children are able to conserve volume during this stage, but some are unable to conserve volume until the formal operational stage. During concrete operations children are able to reverse their thought processes (Muuss, 1996). This means that they can think back to an occurrence and relate it to a current situation which may or may not be abstract. This can be observed when a child is explaining a task of conservation. For example, a child can be given two balls of clay of equal sizes and two glasses of water filled half full. The child can observe the balls being placed into the water so that he/she knows the clay balls will cause the water level to rise to the same height. After the balls are taken out one can be rolled into a sausage, the concrete operational child would be able to explain that if the sausage and the ball were each placed back into their glass, the water levels would again rise to the same levels.

In order to utilize these four elements of logical thinking, children must have developed elementary groupings. Combinativity is the mental ability to combine classes. Reversibility is the ability to reverse the thought process and is an important indicator of development. Children begin to understand that there are different ways to reach the same answer; this understanding is called associativity. Understanding that doing the opposite of one thing undoes the original thing they are able to nullify.

*Formal operational stage.* The final stage in Piaget's theory is formal operations. The largest difference between formal and concrete operations exists in the ability to reason abstractly. Children in the stage of concrete operations can reason about objects where as those in formal operations can reason with thoughts and words alone (Muuss, 1996). The formal operational stage of development may begin around 11 years of age

and is continued throughout adulthood. Formal operations is broken into two substages: IIIA: Almost full formal operations and IIIB: Full formal operations. It is important to note that not all people reach the stage of formal operations. Research suggests that formal education maybe necessary to achieve formal operational theory (Crain, 2005). Piaget suggested that people reached formal operations in their strongest area of interest (Crain, 2005). It could be inferred that people reach formal operations in the area they find most interesting and also choose to study this area more so than others; research does not substantiated which variable (interest or education) comes first and which causes formal operational thought to develop.

During formal operations, individuals are able to use the types of thinking used in previous stages, but in addition they use abstract thinking. Throughout formal operations thought becomes more advanced; this means that mental operations become more abstract, ordered, and logical. Flexibility in thought also increases throughout formal operations. People are no longer governed by reality but can think hypothetically (Muuss, 1996). Deductive reasoning is a central element of formal operational thought (Cowan, 1967). This is the ability to draw valid conclusions from an argument whether or not the argument appears to be valid.

A person could be told, "Mice are bigger than dogs, dogs are bigger than elephants." When asked how mice are related to elephants, a formal operational person would reply that mice are larger than elephants or elephants are smaller than mice.

"The most distinctive property of formal thought is this reversal of direction between reality and possibility; instead of deriving a rudimentary type of theory from the empirical data as is done in concrete inferences, formal thought begins

with a theoretical synthesis implying that certain relations are necessary and thus proceeds in the opposite direction... This type of thinking proceeds from what is possible to what is empirically real” (Inhelder & Piaget, 1958, 241).

New structures, called formal or secondary groupings, are the theoretical foundation of formal logic. Secondary groupings are the cognitive structures which enable adolescents to perform the mental operations characteristic of formal operations. For example, hypothetical-deductive reasoning can now be utilized in order to make inferences and logical conclusions from the general to the specific.

Piaget calls the group of structures that enables the mind to interrelate identity, negation, reciprocity, and correlation the INRC group (Cowan, 1978). “Piaget considered the INRC group as representing the capabilities of most adolescents and normal adults in mathematical-logical or deductive reasoning” (Muuss, 1996, p.162). The “I” stands for identity and is the ability to understand how something can be changed without creating fundamental differences (e.g., the identity property of addition tells that when zero is added to any real number,  $a$ , the result is the number,  $a$ ). Negation or inversion is represented by the letter “N” represents. This is the ability to understand how to undo an original operation (e.g.,  $-5+5=0$ ). The result of an inversion is always an annulment. “R” is for reciprocal or reciprocity. It too undoes an original operation but is different from negation in that reciprocity does not result in an annulment; instead it neutralizes one factor. Reciprocity enables the systematic testing of hypotheses and the manipulation of variables. “C” is for correlative and is the ability to correlate two separate relationships between different situations and understand that if something changes in one relationship, it will also change in the other (e.g., If the relationship

between “A and D” is the same as the relationship between “X” and “Y,” then if “A” causes “D,” “X” causes “Y”).

The combinatorial system of operations results from the development of INRC group. Adolescents are able to combine different variables of this group to make all possible combinations, as the adolescent becomes more advanced; he/she is able to develop a systematic process to achieve each possible combination (Muuss, 1996).

Due to the development of these two structures, adolescents develop formal schemes which allow children to think about relationships between relationships. Adolescents can now conserve volume, judge equivalence, study proportion, and understand issues of relation between time and speed, and think about more complex relationships such as those between individuals and social systems (Cowan, 1978).

Formal operations is broken into two substages. Substage III-A, almost full formal function, is the substage adolescents may enter at 11 or 12 years of age. At 14 or 15 years of age adolescents may gradually enter into substage III-B, full formal function. The difference in these stages is the ability to use systematic abstract thinking. A person performing in substage III-A can think abstractly and reach correct conclusions but may not reach the conclusions in a systematic way. A person performing in substage III-B can also think abstractly; additionally, this person can proceed through the task of solving the problem in a systematically structured way. A person in substage III-A would solve a problem through trial and error where as one in substage III-B would develop a plan before beginning the task at hand. Additionally, people in substage III-B can provide precise proof of their thoughts when questioned about them (Muuss, 1996). The INRC group of operations and the combinatorial system of operations are the two important

theoretical foundations of formal logic (Muuss, 1996). Piaget maintains that without these new structures adolescents would not be able to perform more advanced mental operations than they had during concrete operations (Cowan, 1978).

### *Research of Theory*

Piaget focused his research on the study of how humans acquire and process knowledge; his work opened the doors to a new branch of psychology, genetic epistemology. Genetic epistemology maintains that knowledge results from continuous construction of the mind due to interaction with the environment. Development occurs as individuals move from one stage to the next by the process of formatting new structures of knowledge which did not previously exist (Piaget, 1970). There are many educational implications that can be drawn from his work (e.g., Wadsworth, 1978). Thus, many professionals have devoted much time to researching various parts of his theory. Some researchers attempt to discredit the work of Piaget (e.g., Ennis, 1975), but most research validates his theory of cognitive development.

Piaget's theory suggests that conditional reasoning exists in children operating in the concrete level of development but that formal operational thought may be necessary for proper usage of conditional reasoning in more complex contexts. Kuhn (1977) performed three studies which supported this suggestion. In her first study, Kuhn found 83% and 80% of second and third graders, respectively, were able to make correct assertions using conditional reasoning in simple concrete conversations. Kuhn's study had much higher percentages of success than previous studies assessing the same idea; thus her second study addressed the question of why this should be. She found that the reason children in her study performed better than those in other research studies was due

to the manner of presentation; in her study situations were presented in a concrete, conversational manner. In the third study, children were given a written test which also assessed conditional reasoning. The situations presented in this study were more complex than in the first study.

In this study, little conditional reasoning was seen prior to grade six and no case of complete conditional reasoning was seen before grade eight. Additionally, only two participants in grade four, and half the participants in grade six were able to exhibit the ability to isolate variables at the formal operational level. These findings could be used to support Piaget's assertion that younger students need concrete situations, that they cannot understand abstract concepts such as words on paper; however, it could also be that Kuhn's paper pencil test was poorly written or that students were "helped" when assessed using the conversational method. Therefore, it would be beneficial for other researchers to assess students using both a clinical and a paper-pencil method.

Fischer (1986) performed two studies to determine whether cognitive ability was a predictor of performance in a computer programming course. In the first study she administered *How Is Your Logic?* (a Piagetian based test of cognitive development) with college students addressing their grades in three beginning computer programming classes. This test was used to classify students into a stage of cognitive development. The first study had a sample size of 87 students; 91% of students who received a B+ or higher were in the formal operational stage. No student received a course grade higher than a C+ who was in the concrete operational stage. Of the 87, 29 students were classified as transitional; only two of these received a course grade higher than a B. The correlation between course grade and *How Is Your Logic?* Was statistically significant ( $r$

= .62,  $p \leq 0.05$ ). The study was replicated with a sample of 29 students but there were inconsistencies with grading so the results were not considered valid.

According to Voyat (1983) a study was done in Geneva to determine whether almost illiterate children living in Geneva followed the same developmental sequence with regards to conservation as children living in Europe or America. The study, conducted by Dr. Bovet, found that Genevan children did follow the same sequence, although there was some delay of acquisition (Bovet, 1968; cited in Voyat, 1983). Voyat conducted a study similar to that of Dr. Bovet with a sample of Sioux children. The results were similar to those of the study conducted in Geneva. These two studies support Piaget's theory in regarding to the development of conservation.

Bovet and Voyat's studies also support the notion that cognitive developmental theory is applicable cross-culturally. Crain (2005) cites that several studies have implied that Piaget's stages unfold in the same sequence internationally (Crain, 2005). Eckstein and Shemesh (1992; cited in Muuss, 1996) reported that the pattern of formal operations was similar in the United States and Israel and in many other cross-cultural studies.

Gray (1973) conducted a study which focused on assessing various aspects of four concrete operational groupings and two variations of formal operational characteristics. The participants for the study were 622 nine to fourteen year olds. Two 15-item written tests were given on two consecutive days. The test content assessed various concrete and formal operational structures. The results of the tests confirmed several logical thought models including the formal operations of systematic thinking, making correct and denying incorrect implications, and the concrete operational groupings of additions of asymmetrical relations and addition of asymmetrical relations. In a similar study

conducted by Gray and Hofmann (1976) empirical data supported the “Piagetian postulation that each of the developmental periods is characterized by specific characteristics and a wholistic quality that binds together the various characteristics” (p. 13). Their data specifically supported the idea that combinatorial thought is the basis from which formal operational thought is stemmed.

### *Testing Children for Cognitive Development*

Because his theory provides many educational implications, it would be beneficial for an educator to assess at what level of cognitive development a student was functioning. Throughout his career Piaget developed a group of clinical methods, called tasks, which he presented to children when studying them, generally referred to as a *clinical interview* (Bringuier, 1977). Piaget said that he would suggest a program of experiments to use at the beginning of a year. Then he and his collaborators would develop upon the ideas he presented. The experiments would consist of informal conversations with children; the purpose was finding out new information. In his interview with Bringuier (1977), Piaget went on to explain that he was not testing children because tests are given to determine a level of performance. Instead Piaget (1969) indicated that he and his colleagues were interested in “how the child reasons and how he discovers new tools, so we use direct conversation, informal conversation” (p.24). There are three or four questions Piaget always asked his subjects but otherwise he claimed to explore the mind individually.

In order to determine a child’s level of cognitive development Piaget, or one of his experimenters, would present a child with a series of tasks. When determining the level of development, Piaget was not concerned with whether or not the child produced

the correct answer, but the reasoning behind the answer. Informal conversation was used when assessing all children in order to be able to gauge their thought processes. Using Piaget's clinical method one child is assessed at a time. A child is generally studied throughout the duration of one year; the study ends when the observer stops learning new information (Bringuier, 1977).

According to Bakken (personal communication, December 1, 2005) Piaget suggested a year of practice for the beginning interviewer to be able to master the clinical interview method. Notes should be taken by the task administrator throughout the duration of the interview. The tasks associated with the sensorimotor stage assess secondary circular reactions, systematic exploration, and transition from sensorimotor state or preoperational stage (Crain, 2005). Preconceptual thinking, perspective-taking, and concept permanence are the tasks that should be presented to children who may be in the early preoperational stage. Conservation of number, conservation of continuous quantity, conservation of mass, conservation of weight, and conservation of volume should be assessed for children thought to be in the concrete operational stages. The child thought to be in the formal operational stage should be presented with tasks assessing conservation of volume, combinatorial system of operations, and the ability to use deductive reasoning. The following are two examples of tasks that could be used during a clinical interview to assess a child's level of cognitive functioning (for additional tasks, see Appendix A). The first is a task that assesses concrete operational thinking; the second, formal operational thinking (Wadsworth, 1996).

*Conservation of Mass.* Make two balls of clay that are equal in size. Ask the child if they are equal in size. If the child believes they are not, ask the child to

make adjustments to the balls so that they are equal. After the child confirms that the balls of clay are equal, roll one into the shape of a sausage. Ask the child which has more clay or if they have equal amounts of clay. Next, roll the sausage shaped clay back into a ball. Again, ask the child which has more clay or if the two are equal (Crain, 2005, p.127).

*Combinatorial System of Operations.* Tell the child “The possible life forms are vertebrate, invertebrate, terrestrial, and aquatic. From these life forms, how many possible forms would life take on a new planet?” (Muuss, 1996, p. 162).

#### *Paper-and-Pencil-Tests*

Obviously, it takes an extensive amount of time and training to assess a child’s stage of development using the traditional Piagetian method of clinical interviews. A classroom teacher would be unable to use the clinical interview method as a means of assessing at what cognitive stage each of his/her students was functioning at due to practical constraints. However, it would be beneficial for a teacher to know at what stage of development each student was functioning because the teacher could plan developmentally appropriate lessons. Consequently, several researchers have attempted to develop a group test based on Piagetian principles that could be administered by a teacher to an entire class of students.

Dawson and Rowell (1977) discussed a test which can be used to investigate a student’s understanding of conservation of amount, weight, and volume. The test is said to take twenty minutes to administer and requires the usage of a double pan balance, two one liter beakers, and a piece of plasticine. The article does not say if the test takes twenty minutes to administer to a group or per student. The format of the test requires

that the experimenter manipulate objects while asking the student(s) questions; then the experimenter asks the student(s) to explain the answer. It seems feasible that the test could be administered to a class of students and the students could write their answers on a piece of paper. There are two major drawbacks to this type of administration. First, the method is not that of a typical testing situation so students may be easily distracted, therefore not producing reliable results. Second, the responses would be scored subjectively. Given the limited number of options, this test may still be useful to high school level science teachers. This test could not be used as the sole determinant to classify a student into a stage of cognitive development.

Lawson (1978) conducted a study similar to that of Dawson and Rowell (1977). In his study an experimenter stood in front of a class and manipulated materials to which the participants responded on paper. A purpose of this study included developing a valid instrument with which to measure concrete and formal operational reasoning of secondary school and college age students. The test was found to have “face validity, convergent validity, and factorial validity” (Lawson, 1978, p.21). It is not feasible for classroom teachers to use this test to determine cognitive developmental level of their students because it only assesses concrete and formal operational reasoning abilities and it is not scored objectively. Additionally, if test questions were altered to assess preoperational thought, the method of response by participants may be too difficult for younger students or ESL students.

Walker, Hendrix, and Mertens (1979) discussed a paper-pencil test which included six formal Piagetian tasks which require propositional logic, combinatorial logic, and hypothetical-deductive reasoning. The test was administered to 86 college

students. The statement “explain your answer” was included on most tasks. The responses to this statement were judged on the form of reasoning used and results were compared to check for consistency in judgment. This instrument is not feasible for the classroom teacher because it only assess whether or not a student is in the formal operational stage and because it is not scored objectively. Kenny (1978) conducted a study in which he used the Modified Kenny-Griffiths Test for formal operations to assess his students’ cognitive developmental level. This test was scored objectively; however it could not be used by educators below the high school level because it only assesses the development of formal operations.

Gray (1973) conducted a study which focused on assessing various aspects of four concrete operational groupings and two variations of formal operational characteristics. The participants for the study were 622 nine to fourteen year olds. Two fifteen-item written tests were given on two consecutive days. The test content assessed various concrete and formal operational structures. The results of the tests confirmed several logical thought models including the formal operations of systematic thinking, making correct and denying incorrect implications, and the concrete operational groupings of additions of asymmetrical relations and addition of asymmetrical relations. The test also supports the notion that the existence of logical models can be assessed via paper-and-pencil testing. These findings suggest that the development of a paper-pencil test to detect level of cognitive development is feasible. This test could not be used by a classroom teacher because the test assesses the types of structures a student has, not his level of development.

Gray (1973) constructed a criterion-referenced test intended to be used as a valid and reliable instrument for classroom teachers to use in order to classify a student's stage of cognitive development. The test included 36 multiple choice items, twelve questions corresponded to each of the following: the developmental logic of the pendulum, the developmental logic of the balance, and the developmental logic of combinations of colored and colorless liquids. Half of the questions reflected the logic of concrete operations while the other half reflected the logic of formal operations. Each participant also was assessed using traditional Piagetian tasks. Results found evidence of convergent validity but evidence of discriminate validity was weak. The written portion of this test was scored objectively; however the test format has a major limitation: the written questions were not a written version of the tasks developed by Piaget. Instead they were based on the developer's view of his logic. Additionally, it only assesses the logic of concrete and formal operations; therefore it could not be used by classroom teachers whose students may still be operating at the preoperational stage of development.

A 21-item multiple choice test of Piagetian-tasks was developed by Bakken (1995). This instrument has been used to classify students as operating in one of the following developmental stages: pre-operations, concrete operations: substage one, two, or three or formal operations: substage one or two (Bakken, Thompson, Clark, Johnson, Dwyer, 2001). This is the only paper-pencil test that assesses development at the preoperational level.

The instrument has been used with third, fifth, seventh, ninth, and eleventh grade students who attended public and parochial schools in a metropolitan area. Some students participated in a test-retest format whereas others took the paper-and-pencil test

and were administered the clinical interview one-on-one. Several studies have been conducted in the last few years; however, problems with scoring affected the reliability and validity of the test (e.g. Bird, 2005).

### Summary

Jean Piaget's theory of cognitive development is based on the notion that the interaction of maturation and life experiences is what spurs cognitive development in humans. The four levels of development in Piaget's theory are sensorimotor preoperations, concrete operations, and formal operations; each level has substages. Piaget's theory explains how individuals learn; therefore it has many educational implications. Piaget used time-consuming clinical tasks to assess at what stage of development an individual was operating. Since Piaget, several professionals have attempted to replicate Piaget's clinical tasks via paper-and-pencil testing. So far no one has been able to provide a valid and reliable paper-pencil test that can be used to assess the cognitive levels found in all levels of regular education. Some tests that have been developed are not scored subjectively (Dawson and Rowell, 1977, Lawson, 1978, Walker et al., 1979), many others were not able to assess all stages of development (Gray, 1973, Fischer, 1886, Lawson, 1978). Bakken (1995) developed a paper-pencil test which was used to classify students operating at preoperations, concrete operations, and formal operations; however, the test would be strengthened if the reliability and validity of the instrument were increased.

### *Hypotheses*

There were two hypotheses in this study. First, the test-retest reliability of the instrument would increase due to restructuring of the scoring method and this increase in

reliability would be statistically significantly. Second, the criterion-related validity of the instrument would be increased due to restructuring of the scoring method and this increase in validity would be statistically significant.

## CHAPTER THREE

### THE METHOD

#### *Participants*

The participants included third and fifth grade students from a public dual magnet school which had a high percentage of students who were ESL and came from families with a low SES (N = 133 ). Seventh, ninth, and eleventh grade participants came from three parochial schools that were predominantly white middle class (N = 138). All the students lived in a Midwestern metropolitan area.

Data were been collected at various times throughout the past ten years in which some students were administered the paper-pencil test at three week intervals; other students took the paper-pencil test and were administered the clinical assessment. Data from these studies were used for the purposes of this analysis.

#### *Instrument*

There are 21 multiple-choice items on the objectively scored Piagetian based paper-pencil test (Bakken, et al., 2001). The test was designed to determine at what stage of development a child is operating and can be administered in group settings. Students receive one point for a correct response or zero points for an incorrect response. The test can be used to classify students as preoperational, concrete operational, or formal operational; this test also assesses which substage children are operating in (see Appendix B).

Concrete operational tasks included conservation of number, continuous quantity, length, area, mass, weight, and volume. Right-left hand relationships, classification, and perspective taking were also assessed. Three different substages of concrete operations

were assessed. Substage 1 required 4 of 5 correct responses on conservation of number, continuous length, and area, and on two of the right-left relationship questions. Students who failed to correctly answer 4 out of 5 of these questions were classified as preoperational. Concrete operations, substage 2 required correct answers for substage 1 plus 4 correct answers on mass, weight, right-left relationships, and classification. To be classified as concrete operations substage 3, students must respond correctly to those questions assessed in substage 2 as well as respond correctly to the questions regarding volume and classification (see Table 3).

Formal operational tasks assessed logical reasoning, propositional logic, possibilities, and hypothetical-deductive reasoning. Students can be classified as operating in substage 1 or 2 (using Cowan's [1978] criteria for early and late formal operations). To qualify for formal operations, students need a score of 10 to 12 on the concrete operations questions, plus an 80% rate on each formal operations substage (See Table 3).

Two criterion-related validity studies of the instrument have been conducted (Skar, Bakken, Thompson, Johnson, 1995). A random sample of 40 third grade students from four elementary classrooms participated in the first study. Each child was given the concrete operational Piaget tasks using a clinical interview technique; then the students were administered the multiple choice Piaget test for those items that measured concrete operations. All three substages of the concrete operational stage of the paper-and-pencil test indicated a statistically significant correlation: Concrete operations substage 1 ( $r = .75, p < .01$ ), concrete operations substage 2 ( $r = .69, p < .01$ ), and concrete operations substage 3 ( $r = .69, p < .01$ ) (See Table 1). For the second validation study, a random

sample of 40 fifth grade children from four elementary classrooms were given the Piaget tasks using the clinical interview technique; then the subjects were administered the multiple choice Piaget test for both concrete and formal operations. All three concrete operation substages indicated a significant correlation ( $r = .69, p < .01$ ), and both formal operation substages also yielded a significant correlation ( $r = .54, p < .01$  for the first substage; and  $r = .45, p < .05$  for the second substage) (See Table 1). Thus, the multiple-choice test of Piaget tasks suggests a valid technique of identifying individuals' cognitive thinking, and offers a means of administration to groups at one sitting; however, restructuring of the scoring method may increase the validity of the instrument.

Table 1

*Summary of Validity Coefficients Reported by Skar,, et. al., 1995*

Stage	Grade	Correlation
Concrete, Sub 1	3	0.75 <sup>**</sup>
Concrete, Sub 2	3	0.69 <sup>**</sup>
Concrete, Sub 3	3	0.69 <sup>**</sup>
Concrete	5	0.69 <sup>**</sup>
Formal, Sub 1	5	0.54 <sup>**</sup>
Formal, Sub 2	5	0.45 <sup>*</sup>

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

In order to establish test-retest reliability two studies were done which measured students' responses to the Piaget paper-pencil test at a 3 week interval (Skar, et al., 1995). The first study included 80 seventh grade students from a parochial school in a

metropolitan area. Statistically significant correlations were found for all three substages of concrete operations and both substages of formal operations: Concrete operations substage 1 ( $r = .24, p < .05$ ), concrete operations substage 2 ( $r = .37, p < .01$ ), and concrete operations substage 3 ( $r = .42, p < .01$ ), formal operations substage 1 ( $r = .44, p < .01$ ) and formal operations substage 2 ( $r = .54, p < .01$ ) (See Table 2). The second study assessed a total of 56 students from ninth and eleventh grades at a parochial school in a metropolitan area. Data indicated a significant correlation for all three substages of concrete operations ( $r = .67, p < .01$  concrete operations substage 1;  $r = .63, p < .01$  concrete operations substage 2;  $r = .69, p < .01$  concrete operations substage 3) as well as the two substages for formal operations ( $r = .70, p < .01$  for the first substage and  $r = .61, p < .01$  for the second substage) (See Table 2).

Table 2

*Summary of Reliability Coefficients Reported by Skar, et. al., 1995*

Stage	Grade	Correlation
Concrete, Sub 1	7	0.24**
Concrete, Sub 2	7	0.37**
Concrete, Sub 3	7	0.44**
Formal, Sub 1	7	0.54**
Formal, Sub 2	7	0.45**
Concrete, Sub 1	9-11	0.67**
Concrete, Sub 2	9-11	0.63**
Concrete, Sub 3	9-11	0.69**
Formal, Sub 1	9-11	0.70**

---

Formal Sub 2	9-11	0.61 <sup>**</sup>
--------------	------	--------------------

---

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

Reliability studies indicate that the paper-pencil test is weak to moderately reliable for seventh graders and moderately reliable for high school students.

Restructuring of the scoring method may strengthen the reliability of the instrument.

### *Procedure*

Bird (2005) conducted a study using Bakken's 1995 paper-pencil test. When scoring the protocols, she encountered a problem. An item analysis was completed using a check mark for items missed and correct items were left blank. The design of the assessment should have shown students getting the first questions correct and, as questions grew harder, students should have missed them in order of difficulty. This item analysis illustrated 3 items that were missed by over 85% of the students in the beginning formal stages, but students would then get later-stage formal operation items correct that were successively more difficult. Consequently, items 14, 15, and 16 were eliminated from the new scoring procedure because of extreme item difficulty and stage-like order of development that was missing. The remainder of the items will be included in analyzing this Piagetian paper-pencil assessment. Table 3 illustrates the scoring procedure using both original and revised (Bird, 2005) scoring methods.

Table 3

*Original and Revised Scoring Methods for Piaget Paper-Pencil Test*

Stage	Original Method	Revised Method
Preoperations	Failure to correctly answer four questions out of items 1, 3, 4, 7, 8	N/A
Concrete Sub 1	4 or 5 items correct from 1, 3, 4, 7, 8	N/A
Concrete Sub 2	Above, plus four or five correct items out of 2, 5, 6, 9, 10	N/A
Concrete Sub 3	Above plus correct responses for items 11 and 13	Above plus correct answer for item 11
Formal Sub 1	Above plus 3 or 4 correct items From 12, 14, 15, 17	Above plus correct answers for items 12, 13
Formal Sub 2	Above plus correct answers for items 16, 18, 19, 20, 21	Above plus correct answers for items 18, 19, 20, 21

*Data Analysis*

The researcher re-analyzed existing data for the purposes of this study. In order to establish test-retest reliability a correlational analysis of the new scoring method was to determine the reliability between test and re-test scores. In order to determine criterion-related validity, a correlational analysis between the paper-pencil test using revised scoring and clinical interviews were calculated. A t-test developed to test hypotheses about the equality of two correlation coefficients from dependent samples was used to determine if the revised scoring method was statistically significantly more reliable and more valid than the original scoring method. (Glannapp and Poggio, 1985)

## CHAPTER FOUR

### THE RESULTS

The first hypothesis of this study is that the test-retest reliability of the instrument would increase due to restructuring of the scoring method. A bivariate correlation was conducted to analyze the relationship between test and retest scores on Bakken's paper-pencil instrument; both the original scoring method and the revised scoring method were analyzed (see Tables 4 and 5). A t-test was calculated (Glasnapp & Poggio, 1985) to test for a significant difference in the correlational coefficients between the two scoring methods (see Table 6).

The second hypothesis states that the criterion-related validity of the instrument would be increased due to restructuring of the scoring method. This hypothesis was analyzed using a bivariate correlation to test the statistical significance of the validity between the paper-pencil test and the clinical interview scores; both the original scoring method and the revised scoring method were analyzed. A t-test (Glasnapp & Poggio, 1985) was conducted to test for a significant difference in the correlational coefficients between the two scoring methods

#### *Reliability*

Students were classified into a stage using both the original and revised scoring method for their test and retest scores. Table 4 lists the mean and standard deviation for each test, broken down both by substage and scoring technique.

Table 4

*Means and Standard Deviations for Reliability Coefficients for Third, Fifth, and High School Students on a Piagetian Paper-and-Pencil Test Using Original Scoring and Revised Scoring*

Stage	Test Original Scoring M (SD)	Retest Original Scoring M (SD)	Test Revised Scoring M (SD)	Retest Revised Scoring M (SD)
Preoperations	2.0 (1.03)	2.13 (0.88)	NA	NA
Concrete Sub 1	4.7 (0.46)	4.78 (0.42)	NA	NA
Concrete Sub 2	9.44 (0.79)	9.36 (0.76)	NA	NA
Concrete Sub 3	11.13 (0.79)	11.02 (0.76)	10.23 (0.99)	10.10 (0.98)
Formal Sub 1	13.95 (0.78)	14.0 (0.96)	12.49 (1.61)	12.40 (1.37)
Formal Sub 2	18.0 (1.12)	18.12 (1.29)	15.28 (2.60)	14.72 (2.7)

A bivariate correlation was conducted to test the first hypothesis. Table 5 shows the reliability coefficients for each substage, groups were analyzed according to substage, test, and scoring technique

Table 5

*Reliability Coefficients for Third, Fifth, and High School Students on a Piagetian Paper-and-Pencil Test Using Original Scoring and Revised Scoring*

Stage	N	Original Scoring						Revised Scoring					
		Retest Preop	Retest Conc 1	Retest Conc 2	Retest Conc 3	Retest Form 1	Retest Form 2	Restest Preop	Retest Conc 1	Retest Conc 2	Retest Conc 3	Retest Form 1	Retest Form 2
Test Preop	69	0.35						NA					
Test Concrete Sub 1	81		0.41**						NA				
Test Concrete Sub 2	61			0.67**						NA			
Test Concrete Sub 3	52/61				0.51**						0.79**		
Test Formal Sub1	44/61					0.33*						0.76**	

---

Test				
Formal	40/61		0.26*	0.58**
Sub 2				

---

\* Correlation is significant at the 0.05 level (2-tailed).

\*\* Correlation is significant at the 0.01 level (2-tailed).

The correlational analysis of Bakken's paper-pencil test did not find statistically significant correlation in the preoperational stage of development ( $r = 0.35, p > .05$ ). Every other stage of development was found to have statistically significant correlations regarding both the original and revised scoring methods. Significant correlations for the original scoring methods are as follows: concrete operations substage one ( $r = 0.41, p < .05$ ), concrete operations substage two ( $r = 0.67, p < .05$ ), concrete operations substage three ( $r = 0.51, p < .05$ ), formal operations substage one ( $r = 0.33, p < .05$ ), and formal operations substage two ( $r = 0.33, p < .05$ ).

For the revised scoring, correlations for preoperations, concrete operations substage one, and concrete operations substage two are not analyzed because there were no changes between the original and revised scoring methods for each of these groups.

Correlations for the revised scoring method are as follows: concrete operations substage three ( $r = 0.79, p < .05$ ), formal operations substage one ( $r = 0.76, p < .05$ ), and formal operations substage two ( $r = 0.84, p < .05$ ). Reliability correlations appear to be stronger between test retest scores when the revised scoring procedure is used.

A t-test was conducted to determine whether the differences in the correlations are statistically significant. In order to test two dependent correlations Glasnapp and Poggio (1985) indicate that a t-test after Hotelling is an appropriate statistical analysis.

The following formula was used in the analysis:

$$t = \frac{(r_{13} - r_{23}) \sqrt{[(n-3)(1+r_{12})]}}{\sqrt{[2(1 - r_{13}^2 - r_{23}^2 - r_{12}^2 + 2r_{13}r_{23}r_{12})]}}$$

$r_{13}$  represents the reliability coefficients for the revised scoring method.

$r_{23}$  represents the reliability coefficients for the original scoring method.

$r_{12}$  represents the reliability coefficient between the original and revised scoring method.

See Table 6 for the reliability coefficients between the original and revised scoring methods.

Table 6

*Reliability Coefficients between original and revised scoring method for Third, Fifth, and High School Students on a Piagetian Paper-and-Pencil Test Using Original Scoring and Revised Scoring*

	N	r <sub>23</sub>	r <sub>13</sub>	r <sub>12</sub>	t-value
Concrete Sub3	52	0.51	0.79	0.58	5.02**
Formal Sub 1	44	0.33	0.76	0.78	4.73**
Formal Sub 2	40	0.26	0.58	0.82	5.83**

\* Correlation is significant at the 0.05 level (1-tailed).

\*\* Correlation is significant at the 0.01 level (1-tailed).

r<sub>13</sub> represents the reliability coefficients for the revised scoring method.

r<sub>23</sub> represents the reliability coefficients for the original scoring method.

r<sub>12</sub> represents the reliability coefficient between the original and revised scoring method.

The revised scoring method was statistically significant at the .05 level for each substage analyzed, concrete operations substage three,  $t = 5.02, p < .05$ , formal operations substage one,  $t = 4.73, p < .05$ , formal operations substage two,  $t = 5.83, p < .05$ ; therefore, the researcher's hypothesis that the revised scoring method is more reliable than the original scoring method will be maintained.

### *Validity*

Students were classified into a stage using both the original and revised scoring method for their paper-pencil test scores and their clinical assessments. Table 7 lists the means and standard deviations for each test, broken down both by substage and scoring technique.

Table 7

*Means and Standard Deviations for Validity Coefficients for Third, Fifth, Seventh, and High School Students on a Piagetian Paper-and-Pencil Test and a Piagetian based Clinical Assessment Using Original Scoring and Revised Scoring*

Stage	Test Original Scoring M (SD)	Clinical Original Scoring M (SD)	Test Revised Scoring M (SD)	Clinical Revised Scoring M (SD)
Preoperations	2.16 (0.98)	2.58 (0.65)	NA	NA
Concrete Sub 1	4.72 (0.45)	4.8 (0.41)	NA	NA
Concrete Sub 2	9.40 (0.75)	9.30 (0.78)	NA	NA
Concrete Sub 3	11.07 (0.82)	11.18 (0.83)	10.33 (0.77)	10.27 (0.80)
Formal Sub 1	12.43 (0.50)	13.29 (0.84)	12.04 (0.82)	12.18 (0.81)
Formal Sub 2	16.35 (1.20)	16.68 (1.32)	15.53 (1.15)	12.26 (1.10)

A bivariate correlation was conducted to test the second hypothesis, that the revised scoring method would increase the validity of the instrument. Table 8 shows the reliability coefficients for each substage, groups were analyzed according to substage, test, and scoring technique.

Table 8

*Validity Coefficients for a Piagetian Paper-and-Pencil Test and Piagetian based Clinical Assessment Scores for Third, Fifth, Seventh, and High School Students Using Original Scoring and Revised Scoring*

Stage	N	Original Scoring						Revised Scoring					
		Clinical Preop	Clinical Conc 1	Clinical Conc 2	Clinical Conc 3	Clinical Form 1	Clinical Form 2	Clinical Preop	Clinical Conc 1	Clinical Conc 2	Clinical Conc 3	Clinical Form 1	Clinical Form 2
Test Preop	25	0.22						NA					
Test Concrete Sub 1	103		0.64**						NA				
Test Concrete Sub 2	89			0.62**						NA			
Test Concrete Sub 3	80/89				0.37**						0.58**		
Test Formal	44/74					0.11						0.33**	

---

Sub1				
Test				
Formal	62		0.26	0.47**
Sub 2				

---

\*\* Correlation is significant at the 0.01 level (2-tailed).

The correlational analysis of Bakken's paper-pencil test with the clinical assessment did not find a statistically significant correlation in the preoperational stage of development ( $r = 0.22, p > .05$ ), or for the original scoring method used in defining the formal stage of development substage one ( $r = 0.11, p > .05$ ). Every other stage of development was found to have statistically significant correlations regarding both the original and revised scoring methods. Correlations for the original scoring methods are as follows: concrete operations substage one ( $r = 0.64, p < .05$ ), concrete operations substage two ( $r = 0.62, p < .05$ ), concrete operations substage three ( $r = 0.37, p < .05$ ), and formal operations substage two ( $r = 0.26, p < .05$ ).

For the revised scoring, correlations for preoperations, concrete operations substage one, and concrete operations substage two are not analyzed because there were no changes between the original and revised scoring methods for each of these groups. Correlations for the revised scoring method are as follows: concrete operations substage three ( $r = 0.45, p < .05$ ), formal operations substage one ( $r = 0.33, p < .05$ ), and formal operations substage two ( $r = 0.47, p < .05$ ). Validity correlations appear to be stronger between the paper-pencil test score and clinical assessment scores when the revised scoring procedure is used.

To test for the significance between the correlations of the original scoring and the revised scoring, the t-test that Glasnapp and Poggio (1985) suggested was again used to determine whether the differences in the correlations were statistically significant. See Table 9 for results.

Table 9

*Validity Coefficients between original and revised scoring method for Third, Fifth, Seventh, and High School Students on a Piagetian Paper-and-Pencil Test Using Original Scoring and Revised Scoring*

Stage	N	r <sub>23</sub>	r <sub>13</sub>	r <sub>12</sub>	t-value
Concrete					
Sub3	80	0.37	0.45	0.82	1.31
Formal					
Sub 1	44	0.11	0.33	0.86	5.15 <sup>**</sup>
Formal					
Sub 2	51	0.26	0.47	0.87	5.94 <sup>**</sup>

<sup>\*\*</sup> Correlation is significant at the 0.01 level (1-tailed).

r<sub>13</sub> represents the reliability coefficients for the revised scoring method.

r<sub>23</sub> represents the reliability coefficients for the original scoring method.

r<sub>12</sub> represents the reliability coefficient between the original and revised scoring method.

The revised scoring method was statistically significant at the .05 level for formal operations substage one,  $t = 5.15$ ,  $p < .05$ , formal operations substage two,  $t = 5.94$ ,  $p < .05$ . The difference in the correlations for concrete operations substage three was not statistically significantly significant. The researcher will retain the null hypothesis that the revised scoring method will statistically significantly increase the validity of the instrument because it did in two of the three substages.

## CHAPTER FIVE

### DISCUSSION

This study had two hypotheses. The first was that the test-retest reliability of the instrument would increase due to restructuring of the scoring method and this increase in reliability would be statistically significant. The second was that the criterion-related validity of the instrument would be increased due to restructuring of the scoring method and this increase in validity would be statistically significant.

When the data for reliability was analyzed, all correlations except one were low when using the original scoring method. The revised scoring method made changes regarding the way scoring was done for concrete operations substage three and both substages of formal operations. The correlations were statistically significantly higher for each substage when the revised scoring method was used. The correlations between test-retest were moderately high when the revised scoring method was used.

The revised scoring method eliminated problem test questions. Due to this, the reliability of the instrument increased. This leads the researcher to conclude that several students may have guessed on the eliminated items. The test is much more reliable when items fourteen, fifteen, sixteen, and seventeen are eliminated; therefore, when the test is used in the future, these items should not be used for assessment purposes.

The validity of the instrument was weak to moderate when the original scoring method was used. This caused concern that the instrument may not be assessing what it was designed to assess, cognitive development. When the revised scoring method was used, each substage still had weak to moderate correlations. The t-test found that the revised scoring method was statistically significant at the .05 level for formal operations

substage one and formal operations substage two. The difference in the correlations for concrete operations substage three was not statistically significantly significant. While the revised scoring method increased the validity of the instrument, there is still room to question whether or not the instrument is valid, meaning whether using it would provide the same results as Piaget's clinical tasks. Even though validity correlations are not high, they still exist so there is reason to believe that the test is valid, although there maybe changes that could be made in order to increase it's validity (see *Implications for Future Research* for discussion).

### *Limitations*

The sample used in this study included parochial school students as well as public school students at a dual language magnet that had a high percentage of ESL students. In order to be able to generalize the results to all students, a larger population should be sampled. Subjects in this study should include students from rural, suburban, and metropolitan areas. Subjects should also come from all ethnicities and levels of SES.

### *Implications for Future Research*

In this study, the method of scoring the instrument was examined. Another hypothesis is that problems with reliability and validity occurred due to the way certain test questions were worded. A study should be done in which subjects take the original test and a revised version of the test. The tests could then be scored using both scoring methods. Suggestions for the revised test include adding the sentence "a meal can be composed of one or more foods" to number fourteen, adding "a planet can contain one or more of the life forms" to number fifteen, and changing answer a on number 16 to

“remove one weight from the right side” (Schommer-Aikins, personal communication, 2006).

### *Summary*

Jean Piaget’s theory of cognitive development provides implications for teachers in today’s American classrooms. Teachers who have a basic understanding of Piagetian concepts can structure their lessons so that they are teaching in ways conducive to their student’s cognitive levels. However, in order to do this, teachers have to know what stage of development each of their students is in. A paper-pencil test would enable teachers to find out what stage of development each of their students was functioning in.

Several tests have been developed that attempt to address this need, however none has been published that can be administered to elementary school through high school age students. Bakken’s test has been administered to students ranging through these grades; however her test would be improved if it had stronger reliability and validity coefficients than had been found through previous research. This research reanalyzed Bakken’s data using a restructured scoring method.

The restructuring of the scoring method did increase the reliability and validity of the scoring method. The changes were statistically, significantly different in all areas except for validity, concrete operations substage three. Teachers can be confident when using Bakken’s paper-pencil instrument to classify each of their students as belonging in a particular state of cognitive development. After teachers know what stage of cognitive development their students are functioning in, they can plan and implement lessons that will fit the child’s developmental needs.

## REFERENCES

## REFERENCES

- Bakken, L. (1995). *Fun and challenging puzzles*. Unpublished test. Wichita State University, Wichita, KS.
- Bakken, L., Thompson J., Clark, F., Johnson, N., Dwyer, K. (2001). Making conservationists and classifiers of preoperational fifth-grade children. *The Journal of Educational Research*, 95(1), 56-61.
- Bird, M. (2005). *Linking epistemological beliefs to cognitive development and academic performance*. Unpublished master's thesis, Wichita State University, Wichita, KS.
- Bringuier, J.C. (1977). *Conversations with Jean Piaget*. Chicago, IL: University Press.
- Boden, M. (1979). *Jean Piaget*. New York, NY: The Viking Press.
- College of Counseling, Educational, School Psychology (CESP). (2005). *Testing the Child: Piagetian Activities*. Unpublished work, Wichita State University.
- Cowan, P. (1978). *Piaget: With feeling*. New York: Holt, Rineheart, and Winston.
- Crain, W. (2005). *Theories of development Concepts and applications* (5<sup>th</sup> ed.). Upper Saddle River, NJ: Pearson Education, Inc.
- Dawson, C.J. & Rowell J.A. (1977). Yes. Children are different! The concept of relative density and the conservation of quantities. *The Australian Science Teachers Journal*, 23 (1), 94-97.
- Ennis, R. (1975). Children's ability to handle Piaget's propositional logic. *Review of Educational Research*, 45 (1).
- Fischer, G. (1986). *Computer programming: A formal operational task*. Hiram, OH: Hiram College. (ERIC Document Reproduction Service No. ED275316)
- Funk, J. (2005, February 10) Special report: Grading our schools. *Wichita Eagle*. Retrieved February 3, 2006, from [http://www.kansas.com/multimedia/kansas/archive/2005gos\\_wichita.pdf](http://www.kansas.com/multimedia/kansas/archive/2005gos_wichita.pdf)
- Glasnapp, D. & Poggio, J. (1985). *Essentials of statistical analysis for the behavioral sciences*. Columbus, OH: Charles E. Merrill Publishing Co.
- Gray, W. (1973). *Development of a written test based upon the model of Piaget*. Dayton, OH: University of Dayton. (ERIC Document Reproduction Service No. ED109196)

- Gray, W. (1973). *Development of a Piagetian-based written test: A criterion-referenced approach*. . Dayton, OH: University of Dayton. (ERIC Document Reproduction Service No. ED079325)
- Gray, W. (1973). *The factor structure of concrete and formal operations: A confirmation of Piaget*. Dayton, OH: University of Dayton. (ERIC Document Reproduction Service No. ED115679)
- Gray, W. & Hofmann R. (1976). *Confirmation of the Piagetian logic of exclusion and combinations during concrete and formal operations*. Paper presented at the meeting of the American Educational Research Association, San Fransisco, CA. (ERIC Document Reproduction Service No. ED128686)
- Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Kenny, A. (1978). *The development of a diagnostic instrument based on a Piagetian model*. Nova, FL: Nova University. (ERIC Document Reproduction Service No. ED198876)
- Kuhn, D. (1977). Conditional reasoning in children. *Developmental Psychology*, 13(4), 342-353.
- Lawson, A. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, 15 (1), 11-24.
- Muuss, RE. (1996). *Theories of adolescence* (6<sup>th</sup> ed.). New York: McGraw-Hill.
- Piaget, J. (1967). *Six Psychological Studies*. (Elkind, D. ed) Kingsport, TN: Random House Inc.
- Skar C., Bakken, L., Thompson J., Johnson N. (1995). *Fostering cognitive development: Progressing from Piaget's preoperational stage to concrete operations*. Unpublished manuscript, Wichita State University, Wichita, KS.
- Voyat, G. (1983). *Cognitive development among Sioux children*. New York, NY: Plenum Publishing Co.
- Wadsworth, B. (1977). *Piaget for the classroom teacher*. United Kingdom: Longman Group.
- Wadsworth, B. (1996). *Piaget's theory of cognitive and affective development* (5<sup>th</sup> ed.). New York: Longman.
- Walker, R., Hendrix, J., & Mertens, T. (1979). Written Piagetian task instrument: Its development and its use. *Science Education*, 63(2), 211-220.

## APPENDICES

## Appendix A

*Conservation of number.* One way to assess conservation of number is to make two rows of objects that have the same number of objects. After the child confirms that each row has the same number of objects, the interviewer spaces the row of blocks closer to himself farther apart (not changing the number of blocks in either row). The interviewer then asks the child which row has more blocks or if they have the same amount of blocks. After the child responds the interviewer asks the child to explain why his/her answer is correct. The interviewer will then place the blocks so they are spaced the same as the other row and again ask the child which row has more or if the rows are the same. The interviewer will again ask the child to explain his/her thinking.

*Conservation of Continuous Quantity.* Two glass tumblers of the same size should be filled so that one is  $\frac{2}{3}$  full and the other is  $\frac{1}{3}$  full. After placing the tumblers in front of the child, ask which is more full or if they are the same. If the child believes one to be more full than the other, ask the child to pour water back in forth between the two until they each have equal amounts of water. Next, pour the water from one tumbler into a glass bowl, ask the child whether the tumbler with water in it or the glass bowl has more water or if they are the same.

*Conservation of Mass.* Make two balls of clay that are equal in size. Ask the child if they are equal in size. If the child believes they are not ask the child to make adjustments to the balls so that they are equal. After the child confirms that the balls of clay are equal roll one into the shape of a sausage. Ask the child which has more clay or if they have equal amounts of clay. Next, roll the sausage shaped clay back into a ball. Again, ask the child which has more clay or if the two are equal.

*Conservation of Weight.* Give the child two identical balls of clay, ask the child if the two weigh the same. If the child thinks one weighs more, ask him to make adjustments so that they weigh the same amount. After doing so, roll one ball of clay into a sausage. Ask the child if the sausage or the ball of clay weigh more or if they same amount. Next, roll the sausage back into a ball and ask the child whether the two weight the same or if one weighs more than they other.

*Conservation of Volume.* Fill two glass tumblers half full of water. Ask the child if they have the same amount of water, the child may make adjustments until he/she is confident that they do. Make two balls of clay equal in size, ask the child if they are equal in size and allow the child to make adjustments to them if they are not. Put one ball of clay in each tumbler. Ask the child if the height of the water is the same in each tumbler. If the child says it is not, ask him/her to pour the water back and forth until the amount of water in each is equal. Remove both pieces of clay. Roll one into a sausage. Tell the child you are going to place the ball of clay back into one tumbler and the sausage back into the other. Ask the child which tumbler's water will move up higher or if they will go up the same amount. Ask them to explain their response. Put the clay in the water. If the result was different than the child predicted, ask the child to explain why his/her prediction was incorrect.

*Combinatorial System of Operations.* Tell the child "The possible life forms are vertebrate, invertebrate, terrestrial, and aquatic. From these life forms, how many possible forms would life take on a new planet?" (CESP, 2005, 7)

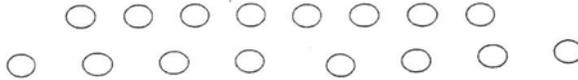
*Deductive Reasoning.* If "A" implies "B" and "B" implies "C" is there a relationship between "A" and "C"? If so, what is the relationship?

## Appendix B

### FUN AND CHALLENGING PUZZLES

**Instructions:** The following are puzzles that we would like you to solve. To the best of your ability, try to find the answer to each puzzle; and then draw a circle around the letter that gives you the answer. If a certain puzzle doesn't make any sense to you, just skip it and go on to the next puzzle.

1. Pretend that these circles are quarters.



Which row has more quarters?

- a. The top row
  - b. The bottom row
  - c. Both rows have the same number.
  - d. You can't tell which row has more.
2. Pretend that these two glasses are  $\frac{2}{3}$  full of coke.



Now, pretend that you take the glass on the left and pour it into this glass:

So now you have these three glasses, one without coke, and two with coke in them:



Which of the two glasses with coke has more coke in it?

- a. The glass on the right has more.
- b. You can't tell which glass has more.
- c. The glass on the left has more.
- d. They both have the same amount.

3. Look at the following two lines, which are the same length:

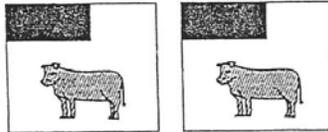


Now, pretend that I move the bottom line so that the two lines look like this:

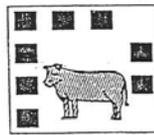


Which line is longer?

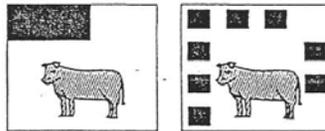
- a. Both lines are the same length.
  - b. The top line is longer.
  - c. You can't tell which line is longer.
  - d. The bottom line is longer.
4. Pretend that the following two squares are two fields with a cow in each field. In the corner of each field is grass for the cows to eat.



Now pretend that you take the grass in the field on the right, cut it up, and move it so it's like this:



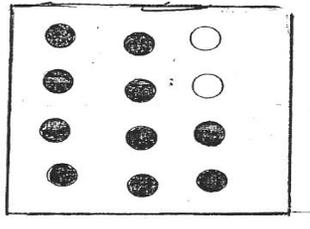
Now you have two fields with cows and grass that look like this:



In which field does the cow have more grass to eat?

- a. You can't tell.
- b. They both have the same amount.
- c. The field on the right has more.
- d. The field on the left has more.

5. Let's pretend that we have this box with twelve plastic beads in it. Some of the beads are black and some of the beads are white.



Are there more black beads or more plastic beads in the box?

- a. There are more black beads.
  - b. There are more plastic beads.
  - c. There are the same number of black beads and plastic beads.
  - d. There are actually more white beads.
6. Assume that I have these two balls of clay:



Not only are they the same size, but they also weigh the same amount. Now I'm going to take the ball on the right and roll it into a sausage so that it looks like this:



So now I have two pieces of clay that look like this:



Which of these pieces of clay weighs more?

- a. They both weigh the same.
- b. The piece on the right weighs more.
- c. The piece on the left weighs more.
- d. You can't tell which piece weighs more.

- 7 - 10. Here we have a pencil, a shoe, and a quarter. Now I'd like to ask you some questions about these three items.



7. First, is the pencil on the right or the left of the shoe?
  - a. The pencil is on the right of the shoe.
  - b. The pencil is on the left of the shoe.
8. Second, is the quarter on the right or the left of the shoe?
  - a. The quarter is on the right of the shoe.
  - b. The quarter is on the left of the shoe.
9. Third, is the shoe on the right or the left of the pencil?
  - a. The shoe is on the right of the pencil.
  - b. The shoe is on the left of the pencil.
10. Fourth, is the shoe on the right or the left of the quarter?
  - a. The shoe is on the right of the quarter.
  - b. The shoe is on the left of the quarter.

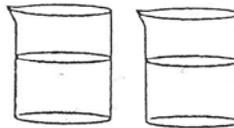
11. Let's take those two balls of clay that are the same size and weigh the same one more time. Now let's pretend that I drop each of them in a glass of water. We can see how the water rises to the same level in each glass.



Now I'm going to take them out of the water and roll this ball on the right into a sausage once more.

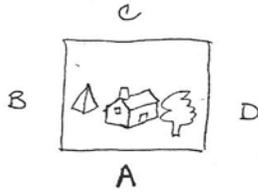


If I put these two pieces of clay back into the water, which piece of clay will make the water rise more?



- a. The clay on the left
  - b. You can't tell which one will make the water rise higher.
  - c. The clay on the right
  - d. They will both make the water rise the same amount.
12. If A is greater than B, and B is greater than C, then C is \_\_\_\_\_ A.
- a. greater than
  - b. the same as
  - c. less than
  - d. similar to

13. Pretend that you are sitting at the table below in position "A."



Now you get up from position "A" and move to position "C." Which of the following would be the correct view of the table if you were sitting in position "C"?



e. None of the above is the correct view.

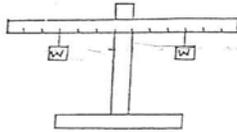
14. Pretend that you have these three foods: ham, cheese, and bread. How many possible different meals can you make from these three foods?

- a. 4
- b. 6
- c. 7
- d. 10
- e. 3

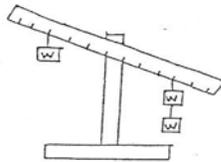
15. Pretend that you just landed on a brand new planet. The life forms that are on this planet are:
- vertebrate
  - invertebrate
  - terrestrial
  - aquatic

From these life forms, how many different possibilities of life forms could there be found on this planet?

- a. 4
  - b. 15
  - c. 18
  - d. 8
  - e. 7
16. Look at this balance beam with a weight on each side of the beam



Now let's add a weight to the right side so it looks like this:



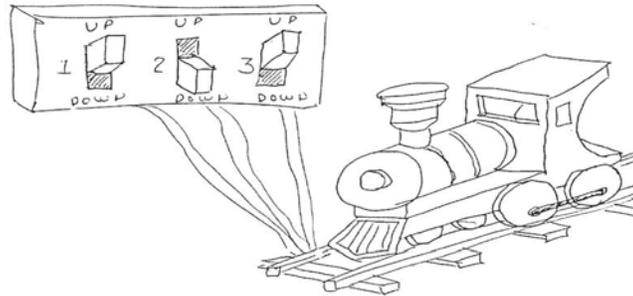
How can the balance beam be brought back in balance again?

- a. Remove the weight.
- b. Add a weight to the left side.
- c. Move the weight on the left farther from the center.
- d. Both a and b, but not c are correct.
- e. Both a and c, but not b are correct.
- f. A, b, and c are all correct.

17. If dogs are bigger than elephants, and elephants are bigger than mice, then dogs are \_\_\_\_\_ mice.
- a. smaller than
  - b. the same size as
  - c. bigger than
  - d. similar to
- 18 - 19. Susan is trying to learn how to play tennis. So she tried several things to improve her game:
- First, she turned her tennis racket a little to the right;  
and she turned her wrist inward;  
and she used type A tennis balls;  
and she hit a good serve over the net.
- Second, she turned her tennis racket a little to the left;  
and she turned her wrist outward;  
and she used type A tennis balls;  
and she hit a poor serve over the net.
- Third, she tried turning her tennis racket a little to the right;  
and she turned her wrist outward;  
and she used type B tennis balls;  
and she hit a poor serve over the net.
- Fourth, she tried turning her tennis racket a little to the left;  
and she turned her wrist inward;  
and she used type B tennis balls;  
and her serve was a good one.
- Last, she tried turning her tennis racket a little to the left;  
and she turned her wrist inward;  
and she used type A tennis balls;
18. Was her serve a good one or a poor one?
- a. a good serve
  - b. a poor serve
19. I decided on the serve because
- a. how she turned her tennis racket was what was important.
  - b. how she turned her wrist was what was important.
  - c. which type tennis balls she used was important.
  - d. all three (turning her tennis racket, her wrist, and the tennis balls) were important.

20 - 21.

Imagine that you are running an electric train that is hooked up to three switches in front of you. Two of the three switches determine how fast the train will go. The way these two important switches are set--down and down, up and down, down and up, or up and up--will determine the speed of the train. The various combinations of positions are given below. Your task is to determine which combination of switches is important and how they work.



<i>Switch 1</i>	<i>Switch 2</i>	<i>Switch 3</i>	<i>Train Goes</i>
Up	Down	Down	Slow
Up	Up	Up	Not at all
Down	Down	Down	Fast
Up	Down	Up	Slow
Down	Up	Up	Slow
Down	Down	Up	Fast
Down	Up	Down	Slow
Up	Up	Down	Not at all

20. Which two switches are important?

- a. 1 and 2
- b. 2 and 3
- c. 1 and 3

21. The way they work is:

- a. When both switches are up, the train goes fast.
- b. When both switches are down, the train goes fast.
- c. When one switch is up and one switch is down, the train goes fast.
- d. When one switch is up and one switch is down, the train does not go at all.
- e. I don't know how they work.