This is to certify that the

thesis entitled

THE DEVELOPMENT, VALIDATION, AND APPLICATION OF AN INSTRUMENT TO ASSESS TEACHERS' UNDERSTANDING OF PHILOSOPHIC ASPECTS OF SCIENTIFIC THEORIES

presented by

Joseph Conrad Cotham

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Secondary Education and Curriculum

Edward L. Smith
Major professor

Date August 9, 1979

THE DEVELOPMENT, VALIDATION, AND APPLICATION OF AN INSTRUMENT TO ASSESS TEACHERS' UNDERSTANDING OF PHILOSOPHIC ASPECTS OF SCIENTIFIC THEORIES

By

Joseph Conrad Cotham

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Secondary Education and Curriculum

1979

THE D
INST
PH

The purp
instrument fo
science that
aspects of s
to describe
preservice
dents, and

The i
significan
sizes its
this inte
tion, ma
science
teaching
concept
tions o
nature
study.

ABSTRACT

THE DEVELOPMENT, VALIDATION, AND APPLICATION OF AN
INSTRUMENT TO ASSESS TEACHERS' UNDERSTANDING OF
PHILOSOPHIC ASPECTS OF SCIENTIFIC THEORIES

By

Joseph Conrad Cotham

The purpose of this study was to develop a reliable and valid instrument for use with elementary and secondary school teachers of science that would assess their conceptions of some philosophic aspects of scientific theories. In addition, the study was intended to describe the results of applying this instrument to samples of preservice elementary teachers, college philosophy of science students, and college chemistry students.

The instrument was organized around an educationally and socially significant interpretation of science, an interpretation that emphasizes its tentative and revisionary characteristics. Understanding this interpretation, in addition to being an important goal of education, may be a significant influence in the successful teaching of science as inquiry. Investigation of the relationship between science teaching and teachers' understanding of the tentative and revisionary conception of science requires a means of assessing teachers' conceptions of science. Thus, concern with teachers' conceptions of the nature of science and their teaching served as justification for this study.

The instrum

(1) it is sensi

aspects of scie

standing of the

science. In r

framework was

scientific th

ontological i

aspect were c

these altern

texts of par

description

This result

five groupi

tion, Opar

and nonthe

The

approache

trait and

struct v

The rela

approach

theories

Th

adminis

30 chem

The instrument was designed to satisfy the following criteria: (1) it is sensitive to alternative conceptions of selected philosophic aspects of scientific theories, and (2) it may be used to infer understanding of the tentative and revisionary conception of the nature of science. In response to these criteria, an instrument-development framework was designed that consisted of five philosophic aspects of scientific theories (i.e., testing, generation, characteristics, ontological implications, and choice). Alternative conceptions of each aspect were described, and items were written to discriminate between these alternative conceptions. Some items were adapted to the contexts of particular scientific theories by prefacing them with a brief description of a scientific theory and episodes drawn from its history. This resulted in an equal distribution of items between the following five groupings: Bohr's theory of the atom, Darwin's theory of evolution, Oparin's theory of abiogenesis, the theory of plate tectonics, and nontheoretical items.

The construct validity of the COST was investigated using two approaches: discrimination between contrasting groups and the multi-trait and multi-method matrix of Campbell and Fiske. Subtest construct validity was supported by the results of these investigations. The relative strength of the validity evidence (obtained from both approaches) is as follows: testing of theories > generation of theories > theory choice ≈ ontological implications of theories.

The final form of the instrument, which contained 50 items, was administered to three groups: 50 elementary education students, 30 chemistry students, and 30 philosophy of science students. Subtest

performances o
of the issues
used as a meas
ception of the
Only 4% o
performed in a
ception of sci
theory generat
formed indeter
subtest altern
the instrument
In contra
science studen
for chemistry
indeterminatel
science and c
a tentative a
In concl
satisfies the
use in assess
nature of sci
with the rela
behaviors.

performances of these groups were analyzed to determine conceptions of the issues embodied in the subtests. Total test performance was used as a measure of understanding the tentative and revisionary conception of the nature of science.

Only 4% of the sample of elementary education students tested performed in a way consistent with the tentative and revisionary conception of science. A majority expressed an induction conception of theory generation. On the remaining subtests, 44% of the sample performed indeterminately, indicating either confusion concerning the subtest alternatives or understanding of conceptions not assessed by the instrument.

In contrast, no more than 27% of the sample of philosophy of science students performed indeterminately on any subtest. Similarly, for chemistry students, no more than 30% of the sample performed indeterminately. Forty-three percent and 27% of the philosophy of science and chemistry students, respectively, performed according to a tentative and revisionary conception of the nature of science.

In conclusion, the instrument that was developed adequately satisfies the development criteria. It is available and suitable for use in assessing teachers' conceptions of particular aspects of the nature of science in research on science teaching which is concerned with the relationship between teachers' conceptions and their teaching behaviors.

To Nancy, whose love and support
provided a much-needed inspiration.

Warm app
man, Ed Smith
essential gui
Bob Floden fo
from his crit
mittee I owe
pragmatic ori
success of th
tion of this
further work.
love and supp
work in persp

ACKNOWLEDGMENTS

Warm appreciation is extended to my friend and committee chairman, Ed Smith, whose sincere concern and helpful criticism provided essential guidance throughout this work. I am especially grateful to Bob Floden for the philosophical and psychometric insight I obtained from his critical comments. And to the remaining members of my committee I owe a special debt of gratitude--to Glen Berkheimer, whose pragmatic orientation spurred me to make corrections essential for the success of this study; and to Andy Timnick, whose ideas on application of this research have provided me with valuable suggestions for further work. Last, but certainly most important, I acknowledge the love and support of my wife, Nancy, whose place in my life puts this work in perspective.

TABLE OF CONTENTS

Chapter

Chapter | Page

Chapter

Chapter | Page

Table

LIST OF TABLES

Table

Table | Page

Figure

LIST OF FIGURES

The purpos

mine the charac

tions of partic

(2) to interpre

drawn from thr

lege philosoph

In 1962

lished the NS

NSTA emphasiz

of elementary

sound science

science is a

Statement on

scientific l

The characte

tion stateme

between fac

knowledge, a

CHAPTER I

INTRODUCTION

Purpose of the Study

The purpose of this study is twofold: (1) to develop and determine the characteristics of an instrument to assess teachers' conceptions of particular aspects of the nature of scientific theories; and (2) to interpret the results of administering the instrument to samples drawn from three populations: preservice elementary teachers, college philosophy of science students, and college chemistry students.

Background

In 1962 the National Science Teachers Association (NSTA) published the NSTA Position on Curriculum Development in Science. The NSTA emphasized understanding the nature of science both as a goal of elementary education and as a requirement for the development of sound science curricula. Emphasis on understanding the nature of science is again evident in the revised document, the NSTA Position Statement on School Science Education for the 70s. In this document, scientific literacy is listed as the major goal of science education. The characterization of scientific literacy presented in this position statement includes behaviors such as identifying the relationship between facts and theory, understanding the tentativeness of scientific knowledge, and understanding the basis for the generation of scientific

knowledge. Th
learning about

Although
goal of scienc
Gale (1966), i
tific literacy
understanding
standing desi
'science is a
also comments
ence found in
tions of scie
from other di
unique" (p. 2
as reflected
many natures.

An inter
is especially
tentative and
of this inter
edge claims i
revision of s
contexts.

Both the
that the scie
eral times du

knowledge. These general characteristics reflect a concern with learning about the nature of science.

Although "understanding the nature of science" is an important goal of science education, its use is ambiguous. Pella, O'Hearn, and Gale (1966), in their analysis of the literature concerned with scientific literacy, concluded that "there is considerable emphasis on the understanding of the nature of science, however, the kinds of understanding desired extended from 'science is a body of knowledge' to 'science is an idea developing activity'" (p. 207). Bridgham (1969) also comments on the variety of interpretations of the nature of science found in the literature of science education. "Authors' conceptions of science range from those that hardly differentiate science from other disciplines to those that make science something observably unique" (p. 26). From these observations it is easy to conclude that, as reflected by the variety of extant interpretations, science has many natures.

An interpretation of the nature of science whose understanding is especially relevant as a goal of science education stresses its tentative and revisionary characteristics. The tentative component of this interpretation emphasizes the inconclusiveness of all knowledge claims in science, and the revisionary component emphasizes the revision of scientific knowledge in response to changing theoretical contexts.

Both the young scientists and the young layman of today will find that the scientific knowledge considered significant will change several times during his life (Robinson, 1968, p. 11). The rapidly changing charac

science educato

dents. In resp

realization tha

additions to kn

of the nature o

revisionary cha

includes an emp

tion of scient

those elements

systems of exp

Robinson

structed and t

construction t

edge that is c

knowledge, the

element of sci

Joseph Schwab

he calls princ

determine what

is given to it

edge construct

> With each
> through us
> embodied a
> which form
> a continui
> inquiry ar
> p. 15).

changing character of scientific knowledge presents a challenge to the science educator in the view of science he will present to his students. In response to this challenge and in order to encourage the realization that important scientific discoveries are never mere additions to knowledge, science educators should teach interpretations of the nature of science that are consistent with its tentative and revisionary characteristics. Attention to these characteristics includes an emphasis on the rule the observer plays in "the construction of scientific knowledge as he observes phenomena and selects those elements of experience which may be constructed into ordered systems of explanation" (Robinson, 1968, p. 124).

Robinson explicates the relationship between the knowledge constructed and the theoretical and metaphysical context in which its construction takes place. As the context changes, so does the knowledge that is constructed. This dynamic characteristic of scientific knowledge, the revisionary nature of science, is possibly the focal element of scientific literacy (Shulman & Tamir, 1973, p. 1101). Joseph Schwab (1962) shows how particular conceptual commitments, which he calls principles of inquiry, structure scientific inquiries. They determine what knowledge is sought, how it is sought, and what meaning is given to it once it has been discovered. And conversely, the knowledge constructed from inquiries suggests new principles of inquiry.

> With each change in conceptual system, the older knowledge gained through use of the older principles sinks into limbo. The facts embodied are salvaged, reordered, and reused, but the knowledge which formerly embodied these facts is replaced. There is, then, a continuing revision of scientific knowledge as principles of inquiry are used, tested thereby, and supplanted (Schwab, 1962, p. 15).

If scientific
ponent of knowledg
scientific knowled

The preceding
ary interpretation
importance in est
cance of this int
ing section.

The justific

1. Understa
the nature of sc
tions. Conseque
goal of science

2. An inst
standing of part
developed for th

a. The
co
th

b. Re
ti
tr

These two argu
discussion.

If scientific knowledge is revisionary, then any particular component of knowledge must be vulnerable to revision. This implies that scientific knowledge is tentative.

The preceding discussion has developed a tentative and revisionary interpretation of the nature of science and has alluded to its importance in establishing goals for science education. The significance of this interpretation to education is addressed in the following section.

Need for the Study

The justification for this study is based on two arguments:

1. Understanding the tentative and revisionary conception of the nature of science has important social and educational implications. Consequently, understanding this conception is an important goal of science education.

2. An instrument that may be used to assess teachers' understanding of particular aspects of the nature of science should be developed for the following reasons:

   a. The goal of understanding the tentative and revisionary conception of the nature of science requires measures that may be used to evaluate its achievement.

   b. Research on the relationship between teachers' conceptions of science and their teaching requires measures that may be used to assess teachers' conceptions.

These two arguments will be addressed successively in the following discussion.

Social and Educ
of Understandin
Revisionary Cha

The social
tative and revi
by considering
Schwab (1962)
for scientists
and the need f
ments are comp
on science and
ally based, mo
or in making
ground, and t
for which sci
pervasive inf
not only more
ciently cogni
mation in ma

However
is the need
ence is incu
society that
nology. Mar
literacy in
Controversi
mentally in

Social and Educational Implications of Understanding the Tentative and Revisionary Characteristics of Science

The social and educational implications of understanding the tentative and revisionary characteristics of science are best approached by considering the educational requirements of our society. Joseph Schwab (1962) has described the most salient requirements as the need for scientists, the requirement for an informed political leadership, and the need for a scientifically literate citizenry. These requirements are compelling because of the increasing reliance of our society on science and technology. As our society becomes more technologically based, more and more people are becoming engaged in activities or in making decisions that require a scientific or technical background, and there is an increasingly wide range of jobs at all levels for which science training is highly useful, if not essential. The pervasive influence of science and technology in our society requires not only more scientists and technicians, but a leadership sufficiently cognizant of science to interpret scientific advice and information in making intelligent decisions affecting the public welfare.

However, perhaps the most significant factor in the social milieu is the need for a scientifically informed public. Knowledge of science is incumbent on citizens who aspire to full participation in a society that is becoming increasingly influenced by science and technology. Many educators have commented on the need for scientific literacy in this society (Hurd, 1958; Johnson, 1962; Shamos, 1963). Controversies over nuclear power, energy conservation, and environmentally induced carcinogenesis are a few examples of socially

significant c
tifically inf
Understa
science has
milieu. Camp
political lea
solution-orie
a certain fle
whose goals a
perspective m
ness of scien
edge that pur
An addit
ception is its
of the scienti
support of the
billion of fed
and technologi
necessity of p
informed publi
encourage thei
Schwab (19
to the need fo
of the tentati
menting on the
important char

significant contemporary issues whose understanding requires a scientifically informed public.

Understanding the tentative and revisionary characteristics of science has implications for meeting the requirements of the social milieu. Campbell (1969) has commented on the desirability of our political leadership developing a problem-oriented rather than solution-oriented perspective. A problem-oriented perspective implies a certain flexibility and openness toward the products of inquiries whose goals are to provide solutions to problems of policy. Such a perspective may be engendered by an understanding of the tentativeness of scientific conclusions and the revisionary nature of the knowledge that purports to provide those conclusions.

An additional consequence of the tentative and revisionary conception is its implication for the public's understanding and support of the scientific enterprise. Public understanding and consequent support of the scientific enterprise is crucial at a time when $30.6 billion of federally administered public monies are spent on scientific and technological research and development (Long & Murray, 1979). The necessity of public support of science requires a scientifically informed public. And yet what must the public know about science to encourage their support of the scientific enterprise?

Schwab (1962) implies that an understanding of science appropriate to the need for an informed public would result from an understanding of the tentative and revisionary characteristics of science. In commenting on the perspective achieved by students who understand these important characteristics of science, Schwab says:

The stud
sarily m
one set
the firs
it was l
may aris
in more
Conseque
for gene
edge (p.

Thus, trust o

essential for

enterprise.

In contr

view that sci

doubt about s

students deve

the changing-

Schwab (1962)

resulting fro

revisionary ch

In conclu

social and edu

tion of scienc

acter. The sc

revisionary c

education.

> The student could understand that to be true does not necessarily mean to be fixed and eternal; that what is said in one set of terms may give way to something else, not because the first was false or has become unfashionable but because it was limited. He could understand that a new formulation may arise and be more desirable because it encompasses more, in more intimate interconnection, than did its predecessors. Consequently, the event of change would no longer be ground for generalized mistrust of the soundness of scientific knowledge (p. 48).

Thus, trust of the soundness of scientific knowledge is construed as essential for public understanding and support of the scientific enterprise.

In contrast to the tentative and revisionary conception, the view that science is a collection of immutable facts can lead to doubt about science and its value. Hillis (1975) speculates that students develop cynicism about science when they are confronted with the changing-knowledge claims of rapidly developing fields of science. Schwab (1962) concurs in this view, seeing cynicism about science resulting from a view of science that neglects its tentative and revisionary character.

In conclusion, evidence has been presented to substantiate the social and educational significance of understanding an interpretation of science which emphasizes its tentative and revisionary character. The social significance of understanding the tentative and revisionary conception underscores its importance as a goal of science education.

Reasons for De
to Determine T
Conceptions of
Aspects of Sci

Measurabi

tentative and
important, the
if attaining a
difference.

The prob
result from u
the nature of
conception we
implications,
to achieving
science, are
tentative and
the specifici
conception.

Understa
understanding
tions of these
standing this
understands a
interpreted i

Joseph S
scientific kn

Reasons for Developing an Instrument to Determine Teachers' (K-12) Conceptions of Particular Philosophic Aspects of Scientific Theories

Measurability. If achievement of the goal of understanding the tentative and revisionary conception of the nature of science is important, then efforts must be made to measure it. This is because, if attaining an educational goal is important, then it must make a difference. Existence of differences suggests measurability.

The problem then becomes one of specifying differences that result from understanding the tentative and revisionary conception of the nature of science. The social implications of understanding this conception were discussed in the previous section. However, these implications, which constitute some differences that may be attributed to achieving the goal of understanding this particular conception of science, are very general. The measurability of understanding the tentative and revisionary conception would be improved by increasing the specificity of the testable implications of understanding this conception.

Understanding the tentative and revisionary conception implies understanding of particular aspects of the nature of science. Descriptions of these aspects provide more specific implications of understanding this conception. The hypothesis, then, that a person understands a tentative and revisionary conception may, therefore, be interpreted in terms of the aspects to be described.

Joseph Schwab (1962) has described the revisionary character of scientific knowledge as an alteration of essential conceptual commitments

by determin

meaning is

theories en

tists in th

ticular asp

revisionary

knowledge i

tific knowl

tific theor

The fol

as categorie

nature of sc

1. Cha
2. Ont
3. Tes
4. Theo
5. Gene

These categor

interpretatio

Importanc

aspects of sci

(K-12) concept

on two justif

and revisiona

important goa

the nature of

his/her teach

commitments. Conceptual commitments structure scientific inquiries by determining what knowledge is sought, how it is sought, and what meaning is given to it once it has been discovered. Scientific theories embody many of the conceptual commitments that guide scientists in their inquiries. Consequently, it is appropriate to use particular aspects of scientific theories in order to express the revisionary nature of scientific knowledge. Because the revision of knowledge implies that it is tentative, the tentative nature of scientific knowledge may also be interpreted in terms of aspects of scientific theories.

The following five aspects of scientific theories are postulated as categories that may be useful in describing conceptions of the nature of science.

1. Characteristics of theories
2. Ontological implications of theories
3. Testing of theories
4. Theory choice
5. Generation of theories

These categories and their relationship to a tentative and revisionary interpretation of the nature of science will be described in Chapter II.

Importance of determining teachers' conceptions of particular aspects of scientific theories. The importance of determining teachers' (K-12) conceptions of particular aspects of scientific theories hinges on two justifications: (1) teachers should understand the tentative and revisionary conception of the nature of science because it is an important goal of science education, and (2) a teacher's conception of the nature of science is a potentially significant factor influencing his/her teaching behavior.

The socia

revisionary co

what is more,

educate future

Because of thi

understand the

ing is expecte

available for

scientific the

tentative and

An additi

of aspects of

is a potential

The significan

that investiga

and their teac

tion is unders

ceptions of th

alternative co

influence scie

measures of te

of science--c

a tentative a

conceptions o

tion of the r

The social significance of understanding the tentative and revisionary conception applies to teachers in their role as citizen. What is more, teachers have a special responsibility in society to educate future scientists, political leaders, and the general public. Because of this responsibility it is especially important that they understand the conception of the nature of science whose understanding is expected of their students. Consequently, means should be available for determining teachers' conceptions of those aspects of scientific theories which may be used to infer understanding of the tentative and revisionary conception of science.

An additional justification for determining teachers' conceptions of aspects of the nature of science is that a teacher's conception is a potentially significant factor influencing teaching behavior. The significance of the tentative and revisionary conception suggests that investigations of the relationship between teachers' conceptions and their teaching should determine the extent to which this conception is understood by teachers. At the same time, alternative conceptions of the specified aspects of scientific theories exist. These alternative conceptions, which will be described in Chapter II, may influence science teaching. It is important, therefore, to develop measures of teachers' conceptions of particular aspects of the nature of science--conceptions of aspects of scientific theories implied by a tentative and revisionary interpretation of science--and alternative conceptions of those aspects which might imply a different interpretation of the nature of science.

The nee

(1) the soci

tative and

its importa

should be d

philosophic

of the goal

of science

ment, (b) i

tative and

used to in

aspects of

The g

Chapter II

ment. Thi

discussion

dation.

instrumen

procedure

described

cedures a

based on

with a su

Summary of the Need for This Study

The need for this study has been addressed from two perspectives: (1) the social and educational implications of understanding the tentative and revisionary conception of the nature of science emphasize its importance as a goal of science education; and (2) an instrument should be developed to determine teachers' conceptions of particular philosophic aspects of scientific theories because: (a) the importance of the goal of understanding the tentative and revisionary conception of science requires that means be developed to measure its achievement, (b) it may be used to infer teachers' understanding of the tentative and revisionary characteristics of science, and (c) it may be used to investigate the relationship between teachers' conceptions of aspects of the nature of science and their teaching.

Organization of the Dissertation

The general organization of the dissertation is as follows: Chapter II contains the theoretical foundation for instrument development. This includes the philosophic basis of the instrument and a discussion of test theory appropriate to its construction and validation. Chapter III consists of a review and critique of related instruments. Chapter IV contains a discussion of instrument-development procedures and results. Performance characteristics of the test are described in Chapter V. Chapter VI contains a discussion of the procedures and results of inferring conceptions of the nature of science based on performance on the instrument. The dissertation concludes with a summary and discussion in Chapter VII.

The the

this chapter

dations, and

foundations

and revisio

sophic aspe

tive and re

ceptions of

validation

investigati

The Tentat
Conception

Expli

nature of

of this co

tentative

ponent and

turn.

CHAPTER II

THEORETICAL FOUNDATION

Introduction

The theoretical foundations of the study, which are included in this chapter, are addressed in two sections: (1) philosophic foundations, and (2) construct validation. The discussion of philosophic foundations will focus on elaboration and explication of the tentative and revisionary conception of the nature of science. Specific philosophic aspects of scientific theories useful in describing the tentative and revisionary conception will be defined, and alternative conceptions of those aspects will be described. The section on construct validation will consist of a discussion of test theory required for investigating the construct validity of the instrument.

Philosophic Foundations

The Tentative and Revisionary Conception of the Nature of Science

Explication of the tentative and revisionary conception of the nature of science requires elaboration of the philosophic implications of this conception. This requirement is addressed by considering the tentative and revisionary conception to consist of a tentative component and a revisionary component, both of which are discussed in turn.

Tentative

results from t

Examination of

the inconclusi

of these argum

Frequentl

tions. Measur

made, and it i

substance have

ioral trait is

is made that t

cies. Both of

science, are i

supposition th

present in mem

impossible to

zation.

An addit

in science is

The hypothes

with cadaver

washing of t

Tentative component. The tentativeness of scientific knowledge results from the inconclusiveness of the arguments used in science. Examination of some patterns of scientific arguments makes explicit the inconclusiveness of the knowledge claims resulting from the use of these arguments.

Frequently in science knowledge claims consist of generalizations. Measurements of samples of a substance's boiling point are made, and it is concluded based on these data that all samples of this substance have a particular boiling point. Or, a particular behavioral trait is observed in a population of organisms, and the claim is made that this trait is characteristic of all members of the species. Both of these generalizations, typical of those developed in science, are inconclusive. There is nothing contradictory in the supposition that all members of a sample have a property that is not present in members of the larger population. Consequently, it is impossible to ascribe certainty to knowledge claims based on generalization.

An additional, related argument used in testing knowledge claims in science is represented in the following scheme:

$$\begin{array}{l} \text{If H, then B} \\ \underline{\text{B}\qquad\qquad} \\ \therefore \text{ H} \end{array}$$

The hypothesis (H) that puerperal fever is caused by contamination with cadaveric matter has as a testable implication (B) that thorough washing of the hands after working with cadavers would result in no infection. A

hypotheses ma

sistent with

of bacteria

this pathoger

cadaveric mat

matter what E

cal proof of

known since a

1973, p. 150)

The type

because their

and found ade

Carl Hempel h

the use of no

fallacy of af

outcome of ev

clusive proof

tial support,

tion for the

does not prov

It is no

used in scien

form:

infection. And yet, demonstration of B does not prove H. Other hypotheses may be consistent with B. For instance, an explanation consistent with modern medical microbiology is that a particular strain of bacteria is the causative agent of puerperal fever. Infection with this pathogen may result from exposure to wound material as well as cadaveric matter. In any case, no matter what H is asserted and no matter what B is derived from H, verification of B provides no logical proof of H. On the contrary, this type of argument is a fallacy known since ancient times as "affirming the consequent" (Ravetz, 1973, p. 150).

The types of arguments cited above, which are nondemonstrative because their premises don't necessitate their conclusions, are used and found adequate by the scientific community. Philosophers such as Carl Hempel have attempted to provide an epistemological rationale for the use of nondemonstrative arguments. However, in referring to the fallacy of affirming the consequent, Hempel (1966) claims that "a favorable outcome of even very extensive and exacting tests cannot provide conclusive proof for a hypothesis, but only more or less strong evidential support, or confirmation" (p. 33). Thus, Hempel's rationalization for the use of a particular type of nondemonstrative argument does not provide any reasons for claiming that it provides certainty.

It is noteworthy that valid demonstrative arguments exist and are used in scientific investigations. Consider arguments of the following form:

This argument

to refute a h

in refuting i

If more than

of most of th

are not all t

is attenuated

> Similarl
> things
> appears
> redefin
> excluded
> lous cas
> of natur

Thus, even t

circumstance

argument doe

scientific h

The pre

arguments we

resulting fr

tific knowle

laws, and hy

this characte

With the qua

"I see a blac

If H, then B

Not B

$\therefore$ Not H

This argument (the *modus tollens*) is deductively valid and may be used to refute a hypothesis. Unfortunately, the strength of this argument in refuting incorrect hypotheses works in only the simplest of cases. If more than one hypothesis is involved in an argument (which is true of most of the arguments in science), we can only conclude that they are not all true. Likewise, the certainty of refuting generalizations is attenuated by the actual practices of scientists.

> Similarly, an assertion of particular properties of a class of things is not simply overthrown when a single contrary instance appears. The original assertion can be defended by a slight redefinition of the class, so that the offending sample is then excluded; or the sample can be dismissed as one of those "anomalous cases" which abound in any detailed study of the workings of nature (Ravetz, 1973, p. 151).

Thus, even though certainty may be ascribed to the *modus tollens*, the circumstances of scientific practice suggest that the use of this argument does not provide absolute assurance concerning the truth of scientific hypotheses.

The previous comments on demonstrative and nondemonstrative arguments were made to emphasize the tentativeness of the knowledge resulting from the use of these arguments. Characterization of scientific knowledge as tentative applies to empirical laws, theoretical laws, and hypotheses. A category of scientific knowledge to which this characterization does not apply is statements of observation. With the qualification of accurate reporting, a statement such as, "I see a black object in front of me," is considered beyond dispute and, therefor

is several st

knowledge--ab

The ten

susceptibili

of scientifi

the revision

following se

<u>Revisio</u>

tific knowle

change in sc

knowledge re

ments to a n

accomplished

knowledge to

type of chan

achieved by

flyby of Jup

sists of act

concluded th

(Morabito, S

example the

tary physics

not previous

knowledge of

edge of plan

and, therefore, conclusive. Much of scientific knowledge, however, is several steps removed from sense data. It is this category of knowledge--abstract, conceptual knowledge--which is tentative.

The tentativeness of much of scientific knowledge implies its susceptibility to change. The characteristics of the changeability of scientific knowledge and their implications for an understanding of the revisionary nature of scientific knowledge are addressed in the following section.

Revisionary component. The revisionary characteristic of scientific knowledge is best approached by distinguishing two types of change in scientific knowledge. The first consists of changes in knowledge resulting from application of a set of conceptual commitments to a new domain of phenomena or to an extent not previously accomplished. The result of this application is the addition of new knowledge to the corpus of scientific knowledge. An example of this type of change is the new understanding of the moons of Jupiter achieved by the Voyager I project. Recent information from the Voyager flyby of Jupiter revealed that Io, one of the moons of Jupiter, consists of active hardrock. In response to this data, astronomers have concluded that Io is the only active, rocky moon in the solar system (Morabito, Synnott, Kupferman, & Collins, 1979, p. 972). In this example the conceptual commitments embodied in the theories of planetary physics and geophysics were applied to an extent and in an area not previously accomplished. The result of this application was new knowledge of a moon of Jupiter and a consequent change in our knowledge of planetary astronomy.

A secon
especially r
conception o
alteration o
tual commitm
the use of o
old and refl

An illu
altered conc
servation of
tual commitm
the principl

> Lavoisi
> apparen
> total m
> he stre
> was an
> the com

Acceptance o
munity both
munity. Mas
gated, syste
and biologic
balance was
of mass cons
tion of the
for conducti
ings.

A second type of change in scientific knowledge, a type that is especially relevant to understanding the tentative and revisionary conception of science, is the revision of knowledge in response to alteration of essential conceptual commitments. Alteration of conceptual commitments results in the replacement of knowledge gained through the use of older principles. The new knowledge is a revision of the old and reflects the insights embodied in the new conceptual systems.

An illustration of the revision of knowledge in response to altered conceptual commitments is the revision of the principle of conservation of mass. Belief in the persistence of matter was a conceptual commitment which was transformed by the scientific community into the principle of conservation of mass.

> Lavoisier demonstrated scientifically that, through all the apparent changes and disappearances of chemical action, the total mass as measured by weight remained unaltered, and thus he strengthened immensely the common-sense view that matter was an ultimate reality, for persistence in time is one of the common-sense marks of reality (Dampier, 1958, pp. 295-96).

Acceptance of the principle of mass conservation by the scientific community both structured and constrained inquiries pursued by this community. Mass balance was expected in the systems that were investigated, systems as diverse as chemical systems created in the laboratory and biological systems studied in the field. Failure to find mass balance was interpreted (in response to the commitment to the principle of mass conservation) as evidence of incomplete and inadequate delineation of the system under study. Subsequent inquiries and strategies for conducting those inquiries were modified in response to such findings.

The p

of natural

relativity

Accor
disti
repre
only

Acceptance

ceptual con

persistence

of this alt

mass conser

lence. Sub

physics) wer

energy conse

Applica

of new knowl

commitments

insights of

Addition

tific knowled

of scientific

tive.

As witn
introdu
rect in
various
science
opment
by a su
1976, p

The principle of conservation of mass functioned as a basic tenet of natural science until the acceptance of Einstein's theory of special relativity in the early part of this century.

> According to the theory of relativity, there is no essential distinction between mass and energy. Energy has mass and mass represents energy. Instead of two conservation laws we have only one, that of mass-energy (Einstein & Infeld, 1938, p. 208).

Acceptance of Einstein's theory of special relativity shifted the conceptual commitment of the scientific community away from belief in the persistence of matter to belief in mass-energy equivalence. The effect of this altered conceptual commitment was revision of the principle of mass conservation to include the implications of mass-energy equivalence. Subsequently, inquiries in particular domains (e.g., nuclear physics) were modified in response to the revised principle of mass-energy conservation.

Application of conceptual commitments may result in the addition of new knowledge to a scientific discipline. Alteration of conceptual commitments results in the revision of old knowledge to reflect the insights of the new conceptual commitments.

Additional insight into the revisionary characteristics of scientific knowledge is obtained by considering two views of the development of scientific knowledge. A prevalent view is that science is cumulative.

> As witnessed by countless references in the forewords and introductions to textbooks, this idea is regarded as a correct interpretation of the historical development of the various disciplines by the representatives of the natural sciences themselves. According to this idea, scientific development consists of a gradual growth of knowledge accompanied by a successive elimination of unscientific ballast (Stegmuller, 1976, p. 137).

Proponents of

be regarded as

theory is see

new theory.

special case

subject to a

velocities of

velocity of l

tic dynamics.

Thomas S

among philoso

accepted only

He contends t

The basi

ferent concep

(1970b) descr

of them, norm

the accepted

lated, and ex

accepted conc

constraints

ence, which

tionary scie

practice of

Kuhn (1

an episode o

Proponents of this view do not deny that occasionally theories come to be regarded as obsolete and are replaced by new ones. But the old theory is seen not to be completely false, but a marginal case of the new theory. For example, Newtonian dynamics has been explained as a special case of relativistic dynamics. Thus, the claim is made, that subject to a number of restrictive conditions (e.g., the relative velocities of the bodies of interest must be small compared to the velocity of light), Newtonian dynamics can be derived from relativistic dynamics.

Thomas S. Kuhn (1970a), who claims he represents a minority view among philosophers of science, asserts that "Einstein's theory can be accepted only with the recognition that Newton's was wrong" (p. 98). He contends that these two theories are fundamentally incompatible.

The basis of Kuhn's contention is found in an essentially different conception of the development of scientific knowledge. Kuhn (1970b) describes two sorts of developmental changes in science. "One of them, normal science, is the generally cumulative process by which the accepted beliefs of a scientific community are fleshed out, articulated, and extended" (p. 250). Normal science, through application of accepted conceptual commitments, produces new knowledge within the constraints imposed by those commitments. In contrast to normal science, which he contends is the more prevalent of the two, is revolutionary science "in which conceptual commitments fundamental to the practice of some scientific specialty must be jettisoned and replaced."

Kuhn (1970a) describes the emergence of relativistic dynamics as an episode of revolutionary science. Even though concepts such as position,

Newtonian

cepts are

that bear

the theor

physical

and commit

In co

the essent

These alte

the succee

altered co

from prece

commitments

knowledge.

ary charact

In the

istics of sc

features of

Alternative

and related

nature of sc

<u>Specific Im</u>
<u>Tentative a</u>
<u>ization of</u>

<u>Aspects</u>

aspects of

position, time, and mass are essential to both relativistic and Newtonian dynamics, "the physical referents of these Einsteinian concepts are by no means identical with those of the Newtonian concepts that bear the same name" (p. 102). The emergence and acceptance of the theory of relativity required that the meaning of fundamental physical concepts had to be revised to agree with the interpretations and commitments of the new theory.

In conclusion, episodes of revolutionary science are times when the essential conceptual commitments of a discipline are altered. These altered conceptual commitments provide a framework within which the succeeding period of normal science occurs. But, in addition, altered conceptual commitments may result in the revision of knowledge from preceding periods of normal science. Thus, altered conceptual commitments exert extensive and pervasive influences on scientific knowledge. These influences underscore the importance of the revisionary characteristics of scientific knowledge.

In the following section the tentative and revisionary characteristics of scientific knowledge will be expressed in terms of specific features of scientific theories. The features will first be defined. Alternative interpretations of the features will then be described and related to the tentative and revisionary interpretation of the nature of science.

Specific Implications of the Tentative and Revisionary Characterization of the Nature of Science

Aspects of scientific theories and their interpretations. Specific aspects of the nature of science may be related to an interpretation of science th

Specificatio

struction. T

ception of t

is tentative

an understan

The rev

cribed above

alterations

ments of a s

used by that

scientific k

claims of sc

ence. Conse

implied by a

and revisio

tion. The

1. Ch
2. On
3. Te
4. Th
5. Ge

These aspec

It is then

be used to

tentative a

of science that emphasizes its tentative and revisionary nature. Specification of these aspects served as a basis for instrument construction. The hypothesis that a person understands a particular conception of the nature of science (such as the conception that science is tentative and revisionary) is, therefore, interpreted in terms of an understanding of the specific aspects to be described.

The revisionary characteristic of scientific knowledge was described above as a revision of scientific knowledge in response to alterations of conceptual commitments. Many of the conceptual commitments of a scientific community are embodied in the scientific theories used by that community. Likewise, the tentative characteristics of scientific knowledge derive from the inconclusiveness of the knowledge claims of science, many of which are included in the theories of science. Consequently, particular aspects of scientific theories, aspects implied by an interpretation of science that emphasizes its tentative and revisionary nature, were chosen as a basis for instrument construction. The following five aspects were used:

1. Characteristics of theories
2. Ontological implications of theories
3. Testing of theories
4. Theory choice
5. Generation of theories

These aspects and their alternative interpretations are described below. It is then shown how particular interpretations of these aspects may be used to infer an understanding of the conception that science is tentative and revisionary.

The inte
from two sou
teachers' co
seven preser
understandin
include info

<u>Charact</u>
around aspec
teachers' co
with this ba
good scienti
1966, p. 75)
theories are
as it is of
scientific t
tions are us
phenomena.
phenomena ex
the theory wa
istics of sc
nations. An
experimental

Empiric
theories. H
exception a
Thus, theori

The interpretations of each aspect that is discussed were obtained from two sources: philosophic literature and probing of elementary teachers' conceptions. Two elementary teachers were interviewed and seven preservice elementary teachers were questioned concerning their understanding of scientific theories. The following interpretations include information from both sources.

Characteristics of theories. Because the instrument is organized around aspects of scientific theories, it is important to determine teachers' conceptions of what theories are. This aspect is concerned with this basic understanding. "The distinctive characteristics of a good scientific theory cannot be stated in very precise terms" (Hempel, 1966, p. 75). In spite of this, some general characteristics of theories are describable. A general property of scientific theories, as it is of any scientific knowledge claim, is empirical import. Any scientific theory must have testable implications. Testable implications are used in verifying the theory by providing explanations of phenomena. In addition, for a theory to be judged good, the range of phenomena explained by the theory must include things not known when the theory was developed (Hempel, 1966, p. 77). The above characteristics of scientific theories relate to their role in providing explanations. An additional role of scientific theories is directing experimental inquiry.

Empirical import is characteristic of hypotheses as well as theories. However, a theory, unlike a hypothesis, is "almost without exception a system of several related statements" (Nagel, 1961, p. 88). Thus, theories manifest a complexity not possessed by hypotheses.

In cont
the "complex
"hypothesis
of theories.
complex, cor
tions of the
educated gue
theories doe
pertains to

<u>Ontolo</u>
has existed
claimed to
status of t
words, is i
ties, event
this issue
the instrum
scientific
primarily
ordering e
ing the ex
tinent or
"the accep
not commit
p. 184).

In contrast to the characteristics just described (representing the "complex view") is a set of characteristics (representing the "hypothesis view") that neglects the complexity and explanatory power of theories. Theories are seen as speculations or guesses rather than complex, conceptual systems. Some preservice teachers' characterizations of theories are "a theory is a hypothesis" and "a theory is an educated guess or explanation of something." This naive view of theories doesn't emphasize the requirement of empirical import that pertains to all scientific theories.

Ontological implications of theories. Considerable controversy has existed in the philosophy of science over what theories may be claimed to assert. This controversy has centered on the ontological status of theoretic entities (Nagel, 1961; Hempel, 1966). In other words, is it appropriate to inquire into the existence of the entities, events, and processes postulated by theories? Two positions on this issue are the instrumentalist and realist views. Advocates of the instrumentalist position emphasize the function of theories in scientific inquiry. An instrumentalist "maintains that theories are primarily logical instruments for organizing our experience and for ordering experimental laws" (Nagel, 1961, p. 118). Questions concerning the existence of the entities postulated by theories are not pertinent or justified from the instrumentalist perspective. Consequently, "the acceptance of theoretical statements when properly understood does not commit us to the existence of theoretical entities" (Brody, 1970, p. 184).

One ar

of theoreti

patible the

argument, c

theories of

century.

light, both

existence'

them, it m

assumed by

native the

Ernes

presents t

Hempel. T

only a tem

developed

This devel

better app

true theor

Accor

theory is

empirical

tain, it i

ties as it

observatio

theory is

One argument used by some instrumentalists against the existence of theoretic entities is based on the use of different and incompatible theories by scientists. Hempel (1966) has discussed this argument, citing the historical example of the wave and corpuscular theories of light before the "crucial experiments" of the nineteenth century. If two theories, such as the two alternative theories of light, both account for the same set of phenomena, then, "if 'real existence' is granted to the theoretical entities assumed by one of them, it must be granted as well to the quite different entities assumed by the other; hence, the entities posited by none of the alternative theories can be held actually to exist" (p. 60).

Ernest Nagel (1961), in describing the realist view of theories, presents the usual reply to the instrumentalist argument discussed by Hempel. This reply asserts that the use of incompatible theories is only a temporary makeshift, to be discarded as soon as a theory is developed which is more comprehensive than either of the previous ones. This development would then be a step in "a series of progressively better approximations to the unattainable but valid ideal of a finally true theory" (p. 144).

According to the realist view (as characterized by Nagel), a theory is literally either true or false. And, even though all empirical knowledge is contingent and not to be characterized as certain, it is as appropriate to assert the existence of theoretic entities as it is to make a similar assertion concerning a matter of observation. "A corollary often drawn from this view is that when a theory is well supported by empirical evidence, the objects ostensibly

postulated by
must be regar
the physical
sticks and s

Testing
two views o
view holds
conclusive
teachers s
and "[a th
there is
of affirm
prove a t

In
a positi
ness of
the ten
chapter
of the
This w
though
favora

hypot
tion
gener

postulated by the theory (e.g., atoms, in the case of atomic theory) must be regarded as possessing a physical reality at least on par with the physical reality commonly ascribed to familiar objects such as sticks and stones" (p. 118).

Testing of theories. This aspect of scientific theories comprises two views of the certainty that may be ascribed to theories. The naive view holds that both scientific hypotheses and theories may be proved conclusively. This view is represented by comments from preservice teachers such as "a theory may be proved to be unquestionably correct" and "[a theory] has been tested and it has reached the point that there is no doubt." Some who hold this view believe that the fallacy of affirming the consequent is a valid argument that may be used to prove a theory or hypothesis.

In contrast to this naive position on the testing of theories is a position (termed the competent view) that emphasizes the inclusiveness of theories and hypotheses. It is represented by the views on the tentativeness of scientific knowledge described earlier in this chapter. The competent view also includes some of the characteristics of the confirmationist approach to hypothesis and theory testing. This widespread approach (Martin, 1972, p. 162) assumes that, even though a hypothesis cannot be proved, it may be supported by the favorable results of testing the implications of the hypothesis.

Generation of theories. This aspect includes two positions on hypothesis and theory generation, induction and invention. The induction conception of theory generation holds that theories are inductive generalizations from empirical data. Included in this position is that

there are "ge
hypotheses or
empirical da

Hempel
"scientific
do not occur
which they
such as "sp
was based o
tra of gase
terms. Th
erate nove

Secor
involving
the curre
different
voltage r
and a cu
between
this hyp
variable
positio
This po
not der
them"
system

there are "generally applicable 'rules of inducation,' by which hypotheses or theories can be mechanically derived or inferred from empirical data" (Hempel, 1966, p. 15).

Hempel has criticized this position from two perspectives. First, "scientific hypotheses and theories are usually couched in terms that do not occur at all in the description of the empirical findings on which they rest" (p. 14). For example, atomic theory contains terms such as "spin," "psi-function," and "electron." Yet, this theory was based on a variety of observations of phenomena (such as the spectra of gases) whose descriptions did not contain these theoretical terms. There is no known set of procedures that may be used to generate novel theoretical concepts from empirical data.

Second, procedures may be described for inferring hypotheses involving simple relationships between variables. For example, if the current of a simple, series circuit has been measured for several different values of voltage, the associated values of current and voltage may be represented by points in a rectangular coordinate system and a curve may be drawn to represent the hypothesized relationship between these two variables. And yet, the procedures which provide this hypothesis presuppose an antecedent hypothesis (relating the two variables) not obtainable by the same procedures. The contrasting position of the theory generation category is the invention conception. This position emphasizes that "scientific theories and hypotheses are not derived from observed facts, but invented in order to account for them" (Hempel, 1966, p. 15). The generation of novel concepts and systems of explanatory constructs cannot be accomplished by any generally

theory dep

The

to mainta

ses or th

The inven

in develo

the contr

scientifi

things, w

investiga

One

of the th

expressed

usefulnes

generatio

vation of

tion is c

view) tha

Theo

acteristi

progresse

a gradual

tion of e

which con

of old kn

generally applicable rules of induction. The transition from data to theory depends on creative imagination.

The induction conception described by Hempel holds that, in order to maintain scientific objectivity, data used in generating hypotheses or theories must be collected without any preconceived ideas. The invention position counters this contention that the data used in developing theories must be collected without preconception. "On the contrary, tentative hypotheses are needed to give direction to a scientific investigation. Such hypotheses determine, among other things, what data should be collected a a given point in a scientific investigation" (Hempel, 1966, p. 13).

One final point remains in describing the alternative positions of the theory generation category. One of the teachers interviewed expressed the belief that how a theory is generated determines its usefulness. This belief was based on an induction conception of theory generation which held that there was an isomorphism between the derivation of a theory from data and how the theory was applied. This position is contrasted with the assertion (included as part of the invention view) that how a theory is generated is irrelevant to its usefulness.

Theory choice. The previous discussion of the revisionary characteristics of science mentioned two conflicting views of how science progresses. The cumulative view holds that science progresses through a gradual growth of knowledge accompanied by the progressive elimination of error. This view conflicts with the revisionary conception, which considers progress in science to be due to a continual revision of old knowledge as it is recast in terms of the new.

These
views of t
consists c

The i
choice cor
tivist pos
vide the b
contrast,
ing theor
idiosyncra
(Kuhn, 19

At th
for theor
terion of
of all ob
theory on
tion as a
judgment
is always
world" (K

If m
then, acc
choose be
teria. B
lished by

These two conceptions of scientific progress imply particular views of theoretical change in science. The aspect of theory choice consists of descriptions of these views.

The issue that is fundamental to the two conceptions of theory choice concerns the basis for theory choice. The traditional, objectivist position asserts the existence of a set of criteria which provide the basis for theory choice. The subjectivist position, in contrast, contends that "the choices scientists make between competing theories depend not only on shared criteria . . . but also on idiosyncratic factors dependent on individual biography and personality" (Kuhn, 1977, p. 329).

At the root of any set of criteria which might provide a basis for theory choice is observation. Subjectivists have called the criterion of observation into question by citing the "theory ladenness of all observational data" (Hanson, 1965, p. 19). The influence of theory on observation, so the subjectivists claim, vitiates observation as an independent standard for evaluating theories. "The act of judgment that leads scientists to reject a previously accepted theory is always based upon more than a comparison of that theory with the world" (Kuhn, 1970a, p. 77).

If more than objective evaluation is involved in theory choice, then, according to the subjectivist position, it is impossible to choose between two rival theories in a domain by using objective criteria. But scientific theories are replaced. If this is not accomplished by comparison with the world, then it is because "a scientific

theory is
to take it

The c
tivist vie
(1972) dis
21).

a. "
t
i
n
g

b. C
t

c. C
b

d. T
t

The la
the ontolog
several exa
(1970a) cla
direction
jectivist
for claimi
is a bette

theory is declared invalid only if an alternate candidate is available to take its place" (Kuhn, 1970a, p. 77).

The objectivist responses to the previously discussed subjectivist views are listed below. They are drawn from Michael Martin's (1972) discussion observation and scientific objectivity" (pp. 116-21).

a. "It is possible to make observational reports in some relatively neutral observational language, i.e., a language that is not free of all theoretical categorization, but that is not in terms of the categories of the theory under investigation" (p. 119).

b. Observation can provide an independent standard for evaluation of scientific theories.

c. Choice between rival theories in a domain may be made on the basis of objective criteria.

d. Theories may be discarded when they conflict with observations.

The last point included in the theory choice aspect concerns the ontological implications of theoretical change. In referring to several examples of theory transition in the physical sciences, Kuhn (1970a) claims that he "can see in their succession no coherent direction of ontological development" (p. 206). Thus, under the subjectivist position is included the contention that there is no basis for claiming that the more recent of a historical pair of theories is a better approximation to the truth. The objectivist response on this poin

is closer

This

particula

will serv

concise fo

Table 1:

Hy
1. Any gue be call
2. A theor able im
3. A theory to expla nomenon.
4. Theories of being things n theory w

Source: Na

this point is that the more recent of a historical pair of theories is closer to the truth (Kuhn, 1970b, p. 265).

This concludes the description of alternative interpretations of particular aspects of scientific theories. These descriptions, which will serve as a basis for instrument construction, are found in a concise form in Tables 1 through 5.

Table 1: Characteristics of Theories

Hypothesis View	Complex View
1. Any guess or speculation may be called a theory.	1. A theory is a system of related statements.
2. A theory needn't have testable implications.	2. Theories have observable implications.
3. A theory is devised in order to explain a puzzling phenomenon.	3. Theories are devised for effectively directing experimental inquiry and for exhibiting connections between matters of observation that would otherwise be regarded as unrelated.
4. Theories needn't be capable of being used to explain things not known when the theory was developed.	4. A theory may be used to explain things not known when it was developed.

Source: Nagel (1961) and Hempel (1966).

Table 2:

1. When a
by empi
and obj
theory
possess

2. A corre
a true
unobser

3. If suff
the evi
postula
ties ma
real.

4. Even th
theories
tists f
only a
ment of
eventua
theory.

Source: Na

Table 2: Ontological Implications of Theories

Realism	Instrumentalism
1. When a theory is well supported by empirical evidence, the events and objects postulated by the theory must be regarded as possessing physical reality.	1. Because theories are primarily logical instruments for organizing experience, theoretic entities cannot be claimed to be physically real on the basis of the acceptance of the theory.
2. A correct scientific theory is a true description of some unobservable reality.	2. The acceptance of theoretical statements when properly understood does not commit us to the existence of theoretical entities.
3. If sufficiently supported by the evidence, theoretically postulated unobservable entities may be claimed to be real.	3. No physical reality may be claimed for the unobservable entities postulated by theories.
4. Even though incompatible theories may be used by scientists for a while, this is only a stage in the development of science that leads eventually to a finally true theory.	4. Because different and incompatible theories may be used by scientists in their research, it is not appropriate to claim that any of them describe what is true.

Source: Nagel (1961) and Hempel (1966).

Table 3: T

1. If a the
tually e
become a
conclusi

2. An argum
form may
theory o

If A is
happen.

B happen

Therefor

Source: Mar

Table 4: G

1. There ar
rules of
hypothes
mechanic
from emp

2. Inductio
be used t
theories

3. The usef
depends
derive t

4. In order
objectiv
erating
must be
preconce

Source: H

Table 3: Testing of Theories

Conclusive	Tentative
1. If a theory is correct, eventually enough evidence will become available to prove it conclusively.	1. Theories may never be proved conclusively.
2. An argument of the following form may be used to prove a theory or hypothesis: If A is correct, then B will happen. B happens. Therefore, A is correct.	2. An argument of the following form may be used to support a theory or hypothesis: If A is correct, then B will happen. B happens. Therefore, A is supported.

Source: Martin (1972).

Table 4: Generation of Theories

Induction	Invention
1. There are generally applicable rules of induction by which hypotheses and theories can be mechanically derived or inferred from empirical data.	1. There are no generally applicable rules of induction by which hypotheses or theories can be mechanically derived or inferred from empirical data.
2. Induction is a method that can be used to derive hypotheses and theories from observed facts.	2. Scientific hypotheses and theories are not derived from observed facts, but invented in order to account for them.
3. The usefulness of a theory depends on the method used to derive the theory from the facts.	3. How a theory is generated is irrelevant to its usefulness.
4. In order to maintain scientific objectivity, data used in generating hypotheses or theories must be collected without any preconceived ideas.	4. Theories and tentative hypotheses determine what data should be collected at a given point in a scientific investigation.

Source: Hempel (1966).

Table 5: T

S

1. It is i
between
domain
objecti

2. Observa
dent st
certain
the inf
observa

3. We have
that th
torical
better
truth.

4. Theorie
other t

5. All des
tions a
retical

Source: K

Aspect Alt
Tentative
acteristic

The r

ception of

section.

scientific

Table 5: Theory Choice

Subjectivist	Objectivist
1. It is impossible to choose between rival theories in a domain solely on the basis of objective criteria.	1. Choice between rival theories in a domain may be made on the basis of objective criteria.
2. Observation is not an independent standard for evaluating certain theories because of the influence of theory on observation.	2. Observation can provide an independent standard for evaluation of scientific theories.
3. We have no basis for claiming that the more recent of a historical pair of theories is a better approximation to the truth.	3. The more recent of a historical pair of theories is closer to the truth.
4. Theories are only replaced by other theories.	4. Theories may be discarded when they conflict with observation.
5. All descriptions of observations are influenced by theoretical preconceptions.	5. Observational reports may be made in some relatively neutral observation language (i.e., a language that is not free of all theoretical categorization, but that is not in terms of the categories of the theory under observation).

Source: Kuhn (1970a, 1970b).

Aspect Alternatives Implied by the Tentative and Revisionary Characteristics of Science

The problem of characterizing the tentative and revisionary conception of science in specific terms was addressed in the preceding section. It was concluded that descriptions of particular aspects of scientific theories were useful in characterizing conceptions of the nature o

are cons

are list

Chara

Ontol
of

Testi

Gener

Theor

Figur

The t

ticated i

fore, tha

sophistica

understand

view of th

The i

does not c

sistent w

knowledge

revision.

when a the

nature of science. Alternative interpretations of these aspects that are consistent with a tentative and revisionary conception of science are listed in Figure 1 and described in the following comments.

Aspect	Alternative Interpretation
Characteristics	Complex view
Ontological Implications of Theories	Instrumentalist
Testing of Theories	Tentative
Generation of Theories	Invention
Theory Choice	Subjectivist

Figure 1. Aspect alternatives consistent with the tentative and revisionary conception of science.

The tentative and revisionary conception of science is a sophisticated interpretation of the nature of science. It is likely, therefore, that teachers who understand this conception will have a sophisticated understanding of theory characteristics. A sophisticated understanding of theory characteristics is represented by the complex view of this aspect.

The instrumentalist contention that the acceptance of theories does not commit us to the existence of theoretic entities is consistent with a view of science that emphasizes that all scientific knowledge is inherently tentative and, therefore, susceptible to revision. In contrast, the realist position on this issue is that when a theory is well supported by the evidence the entities postulated

by that t
ment, a r
sees asse
and open
theory te
a tentati

In t
revisiona
theory an
be compar
theoretic
of scient
tions of
position

The
being der
with a re
are inven
the theor
alternati
of the the
tion conce
tionship b
theories a
dure, then
sistent wi

by that theory may be claimed to exist. On the basis of this commitment, a realist would be less likely to hold a view of science that sees assertions concerning theoretic entities as inherently tentative and open to revision. Similarly, the tentative interpretation of theory testing which disallows assertions of certainty is implied by a tentative and revisionary characterization of science.

In the theory choice aspect, ideas that are consistent with a revisionary and tentative conception of science are the primacy of theory and the rejection of the claim that conflicting theories may be compared with respect to some unchanging, objective reality. The theoretical dependence of observation emphasizes the tentative nature of scientific knowledge. As our theories change, so must our perceptions of the world. The above ideas are represented by the subjective position on theory choice.

The idea that theories are invented to explain data rather than being derived from data using some objective procedure is consistent with a revisionary and tentative conception of science. If theories are invented, then there is no necessary connection between the data the theory is meant to explain and the theory itself. That is, some alternative theory might equally well have been invented. The value of the theory is determined by its usefulness. In contrast, the induction conception of theory generation postulates a much stricter relationship between the theory and the data used in its generation. If theories are derived from data using an established, objective procedure, then this implies a certainty in the derivation that is inconsistent with a tentative and revisionary conception of science. Thus,

the inventi
this concep
The pr
the tentati
alternative
Of the aspe
aspects are
conception
are derived
of the two
is more co
evidence o
this conce
tative and

The p
interpreta
struction.
discrimina
cedures fo
Prior
a few prel
In order t
of the ins
groups: c

the invention interpretation of theory generation is consistent with this conception of science.

The preceding comments have described the relationship between the tentative and revisionary conception of science and particular alternative interpretations of certain aspects of scientific theories. Of the aspects described, the testing of theories and theory choice aspects are most intimately related to the tentative and revisionary conception. It is not claimed that these alternative interpretations are derived from the tentative and revisionary conception. Rather, of the two alternative interpretations of each aspect, one alternative is more consistent with this conception than the other. Consequently, evidence of understanding the alternatives that are consistent with this conception may be construed as evidence of understanding the tentative and revisionary conception of the nature of science.

Construct Validation

The previously described aspects of scientific theories and their interpretations (Tables 1-5) served as a basis for instrument construction. Subtests, organized around each aspect, were devised to discriminate between alternative understandings of each aspect. Procedures for instrument construction are described in Chapter IV.

Prior to addressing construct validation, it is necessary to make a few preliminary comments about the administration of the instrument. In order to provide data used in investigating the construct validity of the instrument, the instrument was administered to three different groups: college chemistry students, philosophy of science students, and elementa

these groups

Interpr

to aspect al

explanatory

vided by inv

and Meehl (1

of people, a

Consequently

ness of inte

structs. T

support of

instrument:

A. Co

1.

2

B. S

3

and elementary education students. Additional information concerning these groups is provided subsequently.

Interpretation of teachers' conceptions of science with respect to aspect alternatives assumes the legitimacy of these aspects as explanatory constructs. Evidence to support this assumption was provided by investigating the construct validity of the aspects. Cronbach and Meehl (1955) describe a construct as "some postulated attribute of people, assumed to be reflected in test performance" (p. 283). Consequently, construct validation was an analysis of the reasonableness of interpreting instrument performance in terms of aspect constructs. The following experimental hypotheses were advanced in support of the validity of the aspect constructs embodied in the instrument:

A. Consistency hypotheses

1. Chemistry students will perform more consistently than elementary education majors as determined by a measure of individual consistency.

2. Philosophy of science students will perform more consistently than elementary education majors as determined by a measure of individual consistency.

B. Subtest differences hypotheses

3. On all subtests, philosophy of science students will perform according to the tentative and revisionary conception more than elementary education majors.

C.

8

The preced

Consistenc

The e

that respo

scientific

with that

istry stud

issues add

directly c

from whom

Their lack

nity to ex

4. For the theory testing subtest, chemistry students will express a competent alternative more than elementary education majors.
5. For the theory generation subtest, chemistry students will express an invention alternative more than elementary education majors.

C. Multi-trait/multi-method hypotheses

6. Requirement 1 of Campbell and Fiske's validation process will be satisfied by all groups.
7. Requirement 2 of Campbell and Fiske's validation process will be satisfied by all groups.
8. Requirement 3 of Campbell and Fiske's validation process will be satisfied by all groups.

The preceding groups of hypotheses will be discussed successively.

Consistency Hypotheses

The expectation that underlies the consistency hypotheses is that respondents who have a particular conception of an aspect of scientific theories will answer instrument items in a way consistent with that conception. Both philosophy of science students and chemistry students have been exposed to educational situations where the issues addressed in this instrument may have been addressed either directly or implicitly. In contrast, the elementary education majors from whom the sample was drawn have had little experience with science. Their lack of experience in science as well as their lack of opportunity to explore issues represented by the instrument aspects would predispos

aspects r

The

consisten

measure o

of item s

a measure

values of

Hull, Jenk

Subtest Di

Due t

science st

standing o

majors. I

read the f

and The St

to the pre

on the iss

well-devel

philosophy

were expect

conception

Likew

cussions o

expected t

predispose this group to an inconsistent understanding of these aspects relative to philosophy of science and chemistry students.

The testing of the consistency hypotheses requires a measure of consistency and statistical tests appropriate to that measure. The measure of performance consistency used in this study was the variance of item scores from each subtest for each respondent. This provided a measure of individual performance consistency. Comparisons of values of consistency across groups were made using a t-test (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975, p. 269).

Subtest Differences Hypotheses

Due to their experience in philosophy of science, philosophy of science students were expected to have a more sophisticated understanding of aspects of scientific theories than elementary education majors. In the courses from which the sample was drawn, students read the following books: Conjectures and Refutations (Karl Popper) and The Structure of Scientific Revolutions (Thomas Kuhn). Exposure to the preceding books in the context of the instructors' emphasis on the issue of scientific objectivity was expected to result in fairly well-developed opinions concerning instrument aspects. Consequently, philosophy of science students more than elementary education majors were expected to perform according to the tentative and revisionary conception of the nature of science.

Likewise, chemistry students because of their exposure to discussions of scientific theories during their chemical education were expected to have fairly sophisticated views of theory testing and theory ge

to discus

two instr

generatio

perform a

than elem

Subt

et al., 1

group's p

significa

<u>Multi-trai</u>
<u>method Hyp</u>

<u>Instr</u>

of Campbel

this secti

variety of

describing

"method" i

The i

consists o

tific theo

of measure

measuremen

which are

Stud

text of p

theory generation aspects. These two aspects appeared more germane to discussions that might occur in science classes than the remaining two instrument aspects. Consequently, on the theory testing and generation of theories subtests, chemistry students were expected to perform according to the tentative and revisionary conception more than elementary education students.

Subtest differences hypotheses were tested using a t-test (Nie et al., 1975, p. 209). Because these hypotheses were in terms of one group's performance exceeding that of another, a one-tailed test of significance was used.

Multi-trait and Multi-method Hypotheses

Instrument methods. The multi-trait and multi-method procedure of Campbell and Fiske (1959) on which the hypotheses described in this section depend requires that test scores are measured using a variety of different methods. It is necessary, therefore, prior to describing this procedure, to explicate the particular meaning given "method" in this context.

The instrument, which will be more fully described in Chapter IV, consists of five subtests based on the particular aspects of scientific theories described previously. Scores for each subtest consist of measurements made with five different methods. The methods of measurement used in this instrument are the different item contexts which are described below.

Students encounter conceptions of scientific theories in the context of particular scientific theories. Thus, even though the

alternative
ence to spe
tations wer
The followi
the atom, (
abiogenesis
cal context
items relat
tion of a t

These
historical
was releva
logical an
theories f
gation of
of the asp

<u>Mult</u>
(1955) st
struct, a
assertion
the same
support o
Campbell
as "a com

The
multi-tra

alternative interpretations in Tables 1 through 5 are without reference to specific scientific theories, items based on these interpretations were adapted to the context of selected scientific theories. The following five items contexts were used: (1) Bohr's theory of the atom, (2) Darwin's theory of evolution, (3) Oparin's theory of abiogenesis, (4) the theory of plate tectonics, and (5) nontheoretical context. Item contexts were created by prefacing each set of items related to a specific context (except #5) by a brief description of a theory and some episodes drawn from its history.

These theories were used to provide item contexts for two reasons: historical information could be found concerning these theories that was relevant to the aspects; and a balance between examples of biological and physical science theories was desired. Use of different theories from the two major fields of natural science allows investigation of the subject-matter dependence of teachers' understanding of the aspects.

Multi-trait and multi-method hypotheses. Cronbach and Meehl (1955) state that "if two tests are presumed to measure the same construct, a correlation between them is predicted" (p. 285). This assertion suggests that significant correlations between scores for the same aspect using different methods would provide evidence in support of construct validity. This position has been elaborated by Campbell and Fiske into what has been described by Magnusson (1967) as "a completely satisfactory validity test" (p. 136).

The procedure of Campbell and Fiske involves the use of a multi-trait and multi-method matrix. Discussion of their procedure will refer

coefficien

five aspec

the aspect

aspects, w

Campb

stitute a

1. T
t
r
T
f

2. T
w
w
w
c
t
i

3. A
t
a
o
t
s
l

4. w
r
c

Requ

different

are valid

construct

Requ

validity

will refer to the partial matrix shown in Figure 2. The correlation coefficients in the matrix will be computed for data obtained in the five aspects using five different methods. The scores for each of the aspects will be correlated with the scores for each of the other aspects, without regard to the method by which they were obtained.

Campbell and Fiske contend that the following requirements constitute a completely satisfactory validation process:

1. The coefficients of correlation between measurements of the same variable (i.e., aspect) with different methods, r_{AB}, r_{AC}, r_{BC}, must be significantly greater than zero. This is the criterion which is normally considered sufficient.

2. The measurements of an apsect must correlate more closely with measurements of the same type which are carried out with another method than with measurements of another type which are carried out with the same method. The validity coefficients, r_{AB}, r_{AC}, r_{BC}, for a certain aspect should thus be greater than the coefficients for the same aspect in the triangles enclosed by solid lines, r_{12}, r_{13}, r_{23}.

3. A validity coefficient for a given aspect must be greater than the correlation between the measurements of this aspect and the measurements of all other aspects with any other method. A validity coefficient should thus be greater than the corresponding coefficients, r_{ab}, r_{ac}, r_{bc}, in the same row and column within the triangle enclosed by dashed lines.

4. Whether the same or different methods are used, the magnitude of the coefficients for the correlation between different aspects should have the same pattern.

Requirement one is concerned with the convergent validity of the different methods of measuring aspect knowledge. If aspect constructs are valid, then different methods of measuring understanding of those constructs should give the same results.

Requirements two and three are concerned with the discriminant validity of different aspects. Thus, measurements of knowledge of different

measuremer

A

B

C

F

No h

has clai

unreliabi

requireme

(p. 137)

Cam

resented

trait th

of metho

different aspects using the same method should correlate less than measurements of knowledge of the same aspect using different methods.

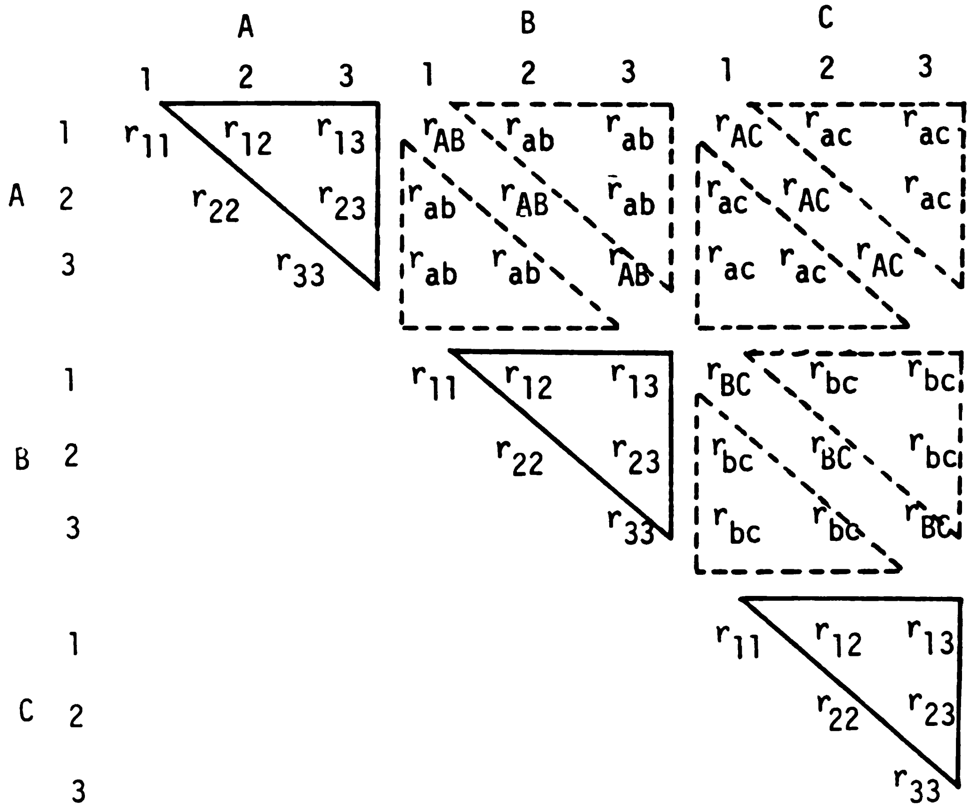


Figure 2. The multi-trait and multi-method matrix.

No hypotheses are based on requirement four. Magnusson (1967) has claimed that "because of the difficulty of judging the effect of unreliability in a matrix of the size we must often deal with, this requirement appears unrealistic and impossible to maintain rigorously" (p. 137).

Campbell and Fiske (1959) have described validity as being "represented in the agreement between two attempts to measure the same trait through maximally different methods" (p. 83). The independence of methods requirement, which is considered by some researchers

(Bouch, M

valid use

in this a

Camp

tion of v

independe

dence for

to a trai

even thou

much in c

of the pr

of the re

It m

Certainly

common.

dence req

differ ma

pletely i

The

cussed in

ception c

particula

tion. Th

(Bouch, Malitz, & Kugle, 1978, p. 127) to be a sine qua non for the valid use of the Campbell and Fiske procedure, is not completely met in this application of their procedure.

Campbell and Fiske (1959) have stated, however, that some evaluation of validity can take place even if the methods are not entirely independent. "In practice, perhaps all that can be hoped for is evidence for relative validity, that is, for common variance specific to a trait, above and beyond shared method variance" (p. 84). Thus, even though the methods of measurement used in the instrument have much in common (all require responses on a written test), application of the procedure of Campbell and Fiske allowed for the determination of the relative validity of the subtest aspects.

It might be added that the term, independent, is ambiguous. Certainly methods that measure a particular trait have something in common. If nothing more, they measure the same trait! The independence requirement should be interpreted as requiring methods that differ maximally with the understanding that they can never be completely independent.

Summary

The theoretical foundations of the proposed instrument were discussed initially. The meaning of the tentative and revisionary conception of science was explicated, followed by a description of particular aspects of scientific theories that related to this conception. The relationship between certain alternative interpretations of these

then desc

procedure

aspects c

Three typ

subtest c

hypothese

the valic

The

understan

for exami

then appl

of these aspects and the tentative and revisionary conception was then described. The second section of this chapter was concerned with procedures for validating the constructs postulated to underlie the aspects of scientific theories that form the basis of the instrument. Three types of hypotheses were described: consistency hypotheses, subtest differences hypotheses, and multi-trait and multi-method hypotheses. Tests of these hypotheses provided evidence to assess the validity of the aspects as explanatory constructs.

The following chapter focuses on extant instruments that measure understanding of aspects of the nature of science. A set of criteria for examining these instruments is described, and the criteria are then applied to a selected set of these instruments.

Deve

of the te

requires

This chap

instrumen

applied t

of the cr

instrumen

nature of

The

two facto

the knowl

knowledge

are boun

of the na

ticular

study was

CHAPTER III

REVIEW OF INSTRUMENTS USED IN ASSESSING UNDERSTANDING OF THE NATURE OF SCIENCE

Introduction

Development of an instrument to assess teachers' understanding of the tentative and revisionary conception of the nature of science requires a set of criteria that may be used in its construction. This chapter will begin with a discussion of criteria appropriate to instrument development. Following this discussion, the criteria are applied to extant instruments as a way of emphasizing the importance of the criteria and, in addition, providing a review and critique of instruments that assess understanding of various aspects of the nature of science.

Instrument-Development Criteria

The criteria used in developing the instrument were based on two factors: (1) the goal of the study, and (2) a consideration of the knowledge domain addressed by the instrument. In review, the knowledge domain of interest is aspects of the nature of science that are bounded by a concern for the tentative and revisionary conception of the nature of science. This has been interpreted to include particular philosophic aspects of scientific theories. The goal of this study was to develop a means of assessing teachers' understanding of the tent

addition

domain.

developme

1.

2.

3.

4.

For

terion of

used to a

tion of s

revisiona

theories.

ing of th

related t

described

claims an

conceptua

The

sensitivi

ence that

criterion

the tentative and revisionary conception of the nature of science in addition to alternative conceptions of aspects of the specified domain. Based on these two factors, the following instrument-development criteria are advocated.

1. The instrument can be used to assess teachers' understanding of the tentative and revisionary conception of science.

2. The instrument is sensitive to multiple conceptions of aspects of the nature of science that are susceptible to a variety of interpretations.

3. The philosophic assumptions that underlie the instrument are explicit.

4. The instrument is organized around sufficiently specific aspects of the knowledge domain so that scores may be interpreted unambiguously.

For the purposes of this study, the instrument-development criterion of first importance is, of course, that the instrument can be used to assess understanding of the tentative and revisionary conception of science. The instrument's sensitivity to the tentative and revisionary conception was based on particular aspects of scientific theories. Other means might be used to assess teachers' understanding of this conception. But, whatever the means used, it must be related to the meaning of the tentative and revisionary conception described previously: the inconclusiveness of scientific knowledge claims and the revision of those claims in response to alterations of conceptual commitments.

The second instrument-development criterion is incorporation of sensitivity to multiple conceptions of aspects of the nature of science that are susceptible to a variety of interpretations. This criterion derives from the conviction that many aspects of the nature of scienc

Lucas, 19

entities,

perspecti

pretation

introduce

instrumen

As d

between t

means of

that seve

among the

of the re

teaching

legitima

conceptio

of conce

issues i

aspects

Exp

an instr

tant bec

light of

complica

a well-a

of science have multiple interpretations (Martin, 1972, p. 153; Lucas, 1975, pp. 481-485). The ontological status of theoretic entities, for example, may be viewed from instrumentalist or realist perspectives. Failure to recognize the several, legitimate interpretations that may be given to aspects of the nature of science introduces the prospect of biased interpretation of performance on instruments that utilize only one conception of science.

As discussed in Chapter I, investigations of the relationship between teachers' conception of science and their teaching require a means of assessing teachers' conceptions. It is reasonable to expect that several misconceptions as well as legitimate conceptions exist among the conceptions possessed by teachers. Consequently, studies of the relationship between teachers' conceptions of science and their teaching should investigate the influence of misconceptions as well as legitimate conceptions. Thus, the criterion of sensitivity to multiple conceptions is interpreted to imply sensitivity to two different types of conceptions: (1) contrasting positions on particular controversial issues in the philosophy of science, and (2) conceptions of particular aspects of the nature of science possessed by teachers.

Explicit specifications of philosophic assumptions that underlie an instrument, the third instrument-development criterion, is important because it allows an interpretation of test performance in the light of those assumptions. Lack of specification of assumptions complicates attempts to explain instrument performance according to a well-articulated view of the nature of science.

A fou
difficulty
knowledge
of instrum
literature
of student
another gr
ing About
Aims of Sc
scales ma
in total

Cons
ment arou
scores ma
Alternati
requisite
the knowl
the fourt
claims.

This
developme
be used
aspects

A fourth instrument-development criterion is suggested by the difficulty of interpreting test scores from instruments that measure knowledge about the nature of science. Aikenhead (1973), in his review of instruments of this type, poses the question: "Throughout this literature, an ambiguity persists: What does it mean for one group of students to have an average TOUS score 4.27 points greater than another group?" (p. 546). TOUS includes three subscales: Understanding About the Scientific Enterprise, The Scientist, and Methods and Aims of Science. The broad range of knowledge covered by these subscales makes it difficult to interpret the significance of differences in total test scores.

Consequently, the fourth criterion is organization of the instrument around specific aspects of the relevant knowledge domain so that scores may be interpreted explicitly in terms of the specified domain. Alternatively, instruments should consist of subtests that embody the requisite specificity. The required domain specificity depends on the knowledge claims made in interpreting test performance. Thus, the fourth criterion should be interpreted in the light of those claims.

This concludes the introduction and discussion of the instrument-development criteria. In the succeeding section these criteria will be used in critically reviewing some instruments that address various aspects of the nature of science.

Review of the Instruments

Table 6 lists the instruments that are reviewed. All of these instruments were developed for use with high school or college students and, therefore, are appropriately administered to elementary or secondary school teachers of science. The instruments are evaluated using the instrument-development criteria. Before considering each instrument, it is noteworthy that none of them can be used to assess teachers' understanding of the tentative and revisionary conception of science. Some of the instruments focus on parts of that conception (e.g., the tentative component), but none address the revisionary component.

Table 6: List of Reviewed Instruments

Instrument	Author
1. Test on Understanding Science (TOUS), Form W	Klopfer and Cooley (1961)
2. Science Process Inventory (SPI), Form D	Welch (1969)
3. Nature of Science Scale (NOSS)	Kimball (1967-68)
4. Views of Science Test (VOST)	Hillis (1975)
5. Science Inventory (SI)	Hungerford and Walding (1974)
6. Wisconsin Inventory of Science Processes (WISP)	Scientific Literacy Research Center (1967)
7. Nature of Scientific Knowledge Scale (NOSKS)	Rubba (1976)
8. Test on the Social Aspects of Science (TSAS)	Korth (1969)
9. Facts About Science Test (FAST)	Stice (1958)

TOUS

This instrument, the most widely used of the ones reviewed, is a four-alternative, 60-item, multiple-choice test. Items are categorized into the three subscales previously mentioned (I--Understanding About the Scientific Enterprise, II--The Scientist; III--Methods and Aims of Science). The topics of subscale III are described in Figure 3.

1. Generalities about scientific methods
2. Tactics and strategy of sciencing
3. Theories and models
4. Aims of science
5. Accumulation and falsification
6. Controversies in science
7. Science and technology
8. Unity and interdependence of the sciences

Figure 3. Subscale III (Methods and Aims of Science).

Because of the focus of this subscale, it is the only one in the instrument that could be used to assess student understanding of the tentative and revisionary characteristics of science. Even within the subscale, there is a considerable range of topics covered. This variety would complicate attempts to interpret subscale scores in terms of the specific themes subsumed by the subscale.

Jungwirth (1974) criticizes the validity of some TOUS items on philosophical and semantic grounds. He claims to "have shown that divergent responses, that is, responses not compatible with the views held by the test authors, may originate in bona fide differences of opinion within the domain of the philosophy of science, and also in misguided linguistic analyses" (p. 210). Jungwirth's detailed item analyses coupled with interviews of respondents revealed that considerable confusion resulted from the use of terms such as "facts," "data," "systematic," and "methodical." Also, in interviewing university professors who had taken TOUS, Jungwirth discovered significant differences in interpretation of key terms that could be attributed to legitimate differences in philosophical viewpoints.

The TOUS is based on a unitary model of the nature of science in spite of controversy over interpretations of key terms used in the instrument. Also, the assumptions on which this model is based are not made explicit in the test manual. Thus, TOUS fails to satisfy criteria #2 and #3. Because of this failure, TOUS is not sensitive to alternative conceptions of aspects of the nature of science that may be important determinants of teachers' behavior. And, failure to specify assumptions precludes interpretation of test performance within the framework of an explicit philosophy of science.

SPI

This instrument is a 135-item forced-choice inventory concerning an understanding of the methods and processes by which scientific knowledge evolves. On the basis of content, the SPI resembles the TOUS subscale III.

Aikenhead's (1972) analysis of the SPI provides some evidence of the difficulty of interpreting scores from instruments that attempt to be comprehensive. He factor analyzed the SPI and found that the factors did not correspond to the original factors predicted by

Welch (Fig

easily int

I. Assu
A.
B.
C.
D.
II. Acti
A.

B.
C.
D.
E.
F.

III. Nat
A.
B.
C.
D.
E.
IV. Et
A.
B.
C.
D.
E.
F.

F

Welch (Figure 4). Also, Aikenhead found that his factors were not easily interpreted.

	Number of Items
I. Assumptions	
A. Reality	
B. Intelligibility	29
C. Consistency	
D. Causality	
II. Activities	
A. Observations	
1. Selection	
2. Infl. past experience	
3. Using instruments	
4. Recording	
5. Describing accurately	
6. Unexpected	
B. Measurement	59
C. Classification	
D. Experimentation	
E. Communication	
F. Mental Processes	
1. Induction	
2. Formulate hypotheses	
3. Deduction	
4. Form. theories, predicting	
5. Many techniques	
III. Nature of Outcomes	
A. Probability	
B. Tentativeness	
C. Theories	
D. Models	37
E. Laws	
IV. Ethics and Goals	
A. Goals and motivation	
B. Objectivity	
C. Anti-authority, skepticism	23
D. Amorality	
E. Repeatability	
F. Parsimony	

Figure 4. Welch's classification of factors in the SPI.

Bates (197
reveal mea
interpreti
attempt to
nature of

The S
ceptions o
tions spec
item from
approximat
an "agree"
tions of s
philosophy
tions unde
discussed

NOSS

The
to measur
sists of
science.
(1) by a
are not

Kim
Conant a

One
on

Bates (1974) also factor analyzed the SPI. His analysis did not reveal meaningful factors. Both studies emphasize the difficulty of interpreting scores obtained from instruments (like the SPI) that attempt to assess understanding of a wide variety of aspects of the nature of science.

The SPI is not intended to discriminate between alternative conceptions of the topics covered by it, nor are the philosophic assumptions specified on which the instrument is based. For example, an item from the SPI states: Science is a series of successively closer approximations to the truth. The scoring key for the SPI indicates an "agree" response for this item. However, the ontological implications of scientific developments are a controversial issue in the philosophy of science (Kuhn, 1970a, p. 206). Nowhere are the assumptions underlying this response or the controversies surrounding them discussed by Welch.

NOSS

The Nature of Science Scale, developed by Kimball, is purported to measure opinions about the nature of science. The instrument consists of 29 statements based on Kimball's model of the nature of science. Students respond to each statement in one of three ways: (1) by agreeing, (2) by disagreeing, or (3) by signifying that they are not sure, do not understand, or feel neutral about the item.

Kimball's model of the nature of science is based on the views of Conant and Bronowski. Martin (1972) has criticized Kimball's model.

> One obvious auxiliary hypothesis is that the model responses on the NOSS reflect an enlightened opinion on the nature of

> science. There is some reason to doubt that this auxiliary hypothesis is true, since at least two of the basic assumptions of the test are dubious: first, there is no one scientific method but only scientific methods, and second, science insists on operational definition (p. 153).

Martin's criticism of the model underlying the NOSS emphasizes that many interpretations of the nature of science are controversial. Thus, even though the philosophic assumptions upon which the NOSS is based are specified, the lack of provision for assessing alternative conceptions in the NOSS biases the interpretation that may be given to performance on the instrument.

VOST

The Views of Science Test was developed by Hillis to measure understanding of the tentativeness of science. It consists of 40 statements that were judged to imply either that science was tentative or absolute. Students expressed the extent of their agreement with the statement, using a Likert scale. Responses were tallied, resulting in a total score that could be interpreted as evidence of students having a tentative, absolute, or mixed conception of science. Even though this instrument can be used to discriminate alternative conceptions, and the assumptions on which the items are based are specified, it has not been divided into subscales. This seems advisable so that interpretations of scores could be used to specify particular understanding or misunderstandings that underlie performance on the instrument. It is noteworthy that the VOST, of all the instruments reviewed, satisfies the instrument-development criteria best. It assesses understanding of the tentativeness of scientific knowledge.

In addition, it is sensitive to multiple conceptions and its assumptions are explicit. However, it is not intended to assess understanding of the revisionary characteristics of scientific knowledge.

<u>SI</u>

The Science Inventory was designed to be sensitive to a variety of different conceptions. It consists of six highly divergent questions (e.g., item 1: Science is...?). Responses to each item are analyzed and grouped into categories. Response categories for item 1 are given in Figure 5. The response analysis and categorization of the Science Inventory has considerable value as a method for identifying the spectrum of conceptions extant in a particular population. However, the generality of the questions used in this inventory precludes its use as an instrument for assessing students' understanding of the tentative and revisionary nature of science.

<u>WISP</u>

The Wisconsin Inventory of Science Processes, constructed by the Scientific Literacy Research Center, consists of 93 statements which a student evaluates as accurate, inaccurate, or not understood. The content of the WISP is almost identical to the SPI. It is not divided into subscales. The breadth of the content covered by this instrument makes it difficult to interpret performance on the instrument in terms of specific aspects of the nature of science. Also, the WISP cannot be used to discriminate alternative conceptions, and the assumptions underlying the instrument are not specified.

1. A resp
a gene
object
ing th

2. A resp
withou
e.g.,
ment.

3. A resp
covery
the er

4. A resp
empiri
tion.

5. Respo
Eithe
of st

6. A res
event

7. Respo

8. A res
categ
that
and s

Figu

NOSKS

The

subscale

school s

Even tho

tions,

knowled

Categories:

1. A response alluding to a combination of processes and product or a generic statement implying this, e.g., a method of studying objects and/or events, empiricism, a logical means of investigating the universe.

2. A response alluding simply to a "study of" objects and/or events without any qualifying statement as to the nature of that study, e.g., exploration of the world, the study of man or the environment.

3. A response alluding to a "discovery phenomenon," e.g., the discovery of living and/or nonliving things or information about the environment.

4. A response alluding to a specific mode of study other than empiricism per se, e.g., hypothesizing, theorizing, prediction.

5. Responses alluding to the knowledge component of science per se. Either as knowledge about objects and/or events or a broad field of study--a content area.

6. A response which tends to equate science with objects and/or events. Science as synonymous with objects and/or events.

7. Respondent did not know and so stated, or item was left blank.

8. A response which was ambiguous and could not be interpreted or categorized logically, e.g., a dialogue with nature; things that one can explore, take apart or add on to; asking questions and seeking answers; a method of learning.

Figure 5. Response categories for item 1: Science is...?

NOSKS

The Nature of Scientific Knowledge Scale is a 48-item, six-subscale, Likert-type research instrument designed to assess high school students' understanding of the nature of scientific knowledge. Even though the instrument is not sensitive to alternative conceptions, it is based on an explicit model of the nature of scientific knowledge. However, of the six model categories only one is appropriate to an und

described

The eight

are gener

understan

ing stude

of the na

Other In

The

of Scien

of this

aspects

tive and

Summary

Th

underst

Table 7

to an understanding of science as tentative and revisionary. It is described in Figure 6.

> Scientific knowledge is never "proven" in an absolute and final sense. It changes over time. The justification process limits scientific knowledge as probable. Beliefs which appear to be good ones at one time may be appraised differently when more evidence is at hand. Previously accepted beliefs should be judged in their historical context.

Figure 6. Developmental category from the NOSKS.

The eight items that constitute the subscale based on this category are general. Scores on this subscale would not reflect specific understandings or misunderstandings that could be used in interpreting student understanding of the tentative and revisionary conception of the nature of science.

Other Instruments

The Facts About Science Test and Test on the Social Aspects of Science were examined and found to be inappropriate to the purposes of this study. Both of these tests are concerned with the social aspects of science and do not address the issues implied by the tentative and revisionary conception of the nature of science.

Summary

The results of the preceding review of instruments that measure understanding of aspects of the nature of science are presented in Table 7. None of the instruments does very well from the perspective

Table 7: Results of Evaluating Instruments Using Instrument-Development Criteria

Instrument	#1--Assessing Tentative and Revisionary Conception	#2--Sensitivity to Multiple Conceptions	#3--Explicit Assumptions	#4--Domain Specificity
TOUS	—	—	—	—
SPI	—	—	—	—
NOSS	—	—	+	—
VOST	P	+	+	—
SI	—	+	NA	—
WISP	—	—	—	—
NOSKS	P	—	+	—

Legend: — = Criterion not satisfied

+ = Criterion satisfied

P = Criterion partially satisfied

NA = Not applicable because of test structure

of criteri
teachers'
of the na
developme
Chapter I

of criteria intended for use in developing an instrument to assess teachers' understanding of the tentative and revisionary conception of the nature of science. Application of these criteria to the develop- development of the instrument for the current study is addressed in Chapter IV.

CHAPTER IV

INSTRUMENT DEVELOPMENT

Introduction

The purpose of this study was the development, application, and evaluation of an instrument to measure teachers' understanding of the tentative and revisionary conception of the nature of science. This chapter includes descriptions of procedures used in development of the instrument and consists of the following: (1) instrument-development procedures, (2) item-selection results, and (3) chapter conclusion.

Instrument-Development Procedures

Instrument Characteristics

Criteria. The instrument-development criteria described in Chapter III were used in planning the construction of the instrument. The first criterion, the instrument may be used to assess understanding of the tentative and revisionary conception of science, was central to the development of the proposed instrument. Particular aspects of scientific theories (discussed in Chapter II), which relate to the tentative and revisionary conception, served as the framework for instrument construction. Organization of the instrument around these aspects makes it useful in assessing understanding of the tentative and revisionary conceptions of science. Because the instrument is organized around aspects of scientific theories, it is called the Conceptions of Scientific theories Test (COST).

The second criterion, sensitivity to multiple conceptions of aspects of the nature of science, was incorporated into the COST by writing items that discriminate between alternative conceptions of each aspect. These alternative conceptions (described in Tables 1-5 of Chapter II) consist of two different interpretations of each aspect. The number of interpretations is two and was limited to this number because only two major positions on certain philosophical issues were found (e.g., ontological implications of theories). In addition, the item format that was used (discussed subsequently) is intended for a dichotomous choice. Consequently, use of this format required that only two interpretations be assessed by each item.

The COST is divided into subtests, each subtest corresponding to a particular aspect of scientific theories. As mentioned, the distinguishing features of the alternative conceptions of each aspect were described in Chapter II. These COST characteristics, subtest specificity and explicit description of subtest alternative, satisfy instrument-development criteria #3 and #4. Criterion #3, specification of philosophic assumptions underlying the instrument, is satisfied by explicit characterization of the alternative conceptions of each aspect which served as the basis for item construction. Criterion #4, sufficient specificity of the knowledge domain to allow relatively precise interpretation of instrument scores, is satisfied by the specificity of the descriptions of the subtest alternatives.

This speci

scores in

<u>Conte</u>

theories i

though the

erences to

natives we

theoretica

items by

drawn fro

the descr

is an exa

in the CO

> In
> the
> atm
> dif
> thi
> can
> the
> the
> mor
> for
>
> Sin
> dor
> tha
> Op
> to
> co
> ar
>
> In

theoret

This specificity allows for an unambiguous interpretation of subtest scores in terms of the subtest alternatives.

Contexts. Teachers encounter philosophic aspects of scientific theories in the context of particular scientific theories. Thus, even though the descriptions of alternative conceptions are without references to specific scientific theories, items based on these alternatives were couched in terms of particular scientific theories. The theoretical context of the items was created by prefacing each set of items by a brief description of a scientific theory and some episodes drawn from its history. Items that follow that description refer to the description and make use of terms included in it. The following is an example of a description used in creating a theoretical context in the COST.

Oparin's Theory of Abiogenesis

> In 1938, a Russian bio-chemist, A. I. Oparin, proposed a theory to explain the origin of life. He argued that the atmosphere of the earth before the origin of life was very different from what it is today. Under these conditions of this early atmosphere, Oparin claimed that simple molecules came together to form more complex organic substances that are the constituents of living systems. Eventually, according to the theory, the organic substances combined together to form more and more complex substances, until a living structure was formed.
>
> Since Oparin developed his theory many experiments have been done to test it. In 1953, Stanley Miller published a paper that described his attempts to test some of the claims of Oparin's theory. Miller simulated conditions that were thought to duplicate those of the earth's early atmosphere. Under these conditions he was able to produce many complex substances that are constituents of living organisms.

In order to facilitate investigation of the effect different theoretical contexts might have on understanding aspects of scientific theories, the theories used in the COST were chosen from the fields of both biological and physical science. An additional reason for choosing the theories used in the COST was the availability of information about historical episodes relevant to those theories that could be used to exemplify the aspects of scientific theories upon which the COST was based. The following four theoretical contexts were used: (1) Bohr's theory of the atom, (2) Darwin's theory of evolution, (3) Oparin's theory of abiogenesis, and (4) the theory of plate tectonics. A fifth context used in the COST did not refer to a specific scientific theory. Items in this context were not prefaced by a description of a scientific theory and were in terms of general characteristics of scientific theories.

Item format. Each item was written as a statement which embodied some feature of the theory aspect alternatives described in Tables 1 through 5 of Chapter II. Items were scaled using a modified Likert scale [(1) strongly agree, (2) agree, (3) disagree, and (4) strongly disagree]. A modified Likert scale is suitable for COST items because they were constructed to discriminate between two conceptions of the aspect represented by each item. In contrast to a dichotomous scale, a modified Likert scale provides information concerning the conviction of response. The "undecided" option of the Likert scale was excluded in order to force respondents to choose one or another aspect alternative. It was assumed that ambiguously worded items that may have made choice of the "undecided" option appropriate were excluded in the item-selection process.

<u>Student information items</u>. Prospective school teachers have had varying amounts of experience with the theories that provide the contexts for the COST items. It is conceivable that the depth of understanding of a scientific theory possessed by a teacher may influence the particular conception of a theory aspect held by that teacher. Thus, it is desirable to include a set of items in the instrument that could be used to evaluate the teachers' knowledge of the theoretical contexts of the instrument as a way of exploring and providing possible explanations of the subject-matter dependence of teachers' conceptions.

Personal bias is another factor that may affect teachers' conceptions of scientific theories. Personal bias is defined as influences, other than subject-matter knowledge, that may affect teachers' performance on the COST. For example, particular scientific theories (e.g., Darwin's theory of evolution) conflict with some systems of religious belief. It is important, therefore, for the COST to include a set of items that could be used to assess the extent of personal bias toward the theoretical contexts of the instrument.

The student information items have a multiple-choice format. Respondents are asked to rate their knowledge on each theory using the following options: (1) mastery, (2) highly competent, (3) somewhat competent, (4) slightly competent, and (5) no knowledge. Personal bias items ask respondents to evaluate the influence of personal conviction of responses to items associated with particular theoretical contexts. Extent of influence is to be rated as (1) complete, (2) strong, (3) moderate, (4) weak, or (5) none.

Item-Selection Procedures

Item-selection procedures consisted of generation of a pool of items, administration of the items to a sample of prospective elementary school teachers, and selection of the best items from that pool. The criteria of item selection were: (1) the integrity of the subtest as measured by its reliability, and (2) the relationship between each item and the subtest score.

Eighty items were written and divided into two sets of 40 to obtain two forms of the COST (pilot forms A and B). Each pilot form consisted of five subsets of eight items. Each subset, which represents a subtest of the COST, consists of eight items based on one aspect of scientific theories. Items within each subtest are found in three different contexts. The organization of the items in each pilot form of the COST is represented in Table 8.

Pilot forms A and B (found in Appendices A and B) were administered to 56 college physical science students during the summer of 1978. These students were primarily elementary education majors. Twenty-nine students took form A, and 27 took form B.

Cronbach's alpha, a measure of reliability, was determined for each subtest. Alpha is the mean of all possible split-half coefficients. In addition, it estimates the proportion of test variance attributable to common factors among the items (Cronbach, 1951). Values of alpha for the subtests were determined and used as estimates of the common-factor concentration of each subtest.

Pearson product-moment correlation coefficients were determined for each item and the total score for the subtest (with that item excluded) to which the item belonged. The item with the lowest correlation coefficient in each cell (subtest x TC_n --refer to Table 8) was deleted because of its relatively weak relationship to the subtest.

Table 8: Structure of COST Pilot Forms

Subtest	Theoretical Context			
	TC_1	TC_2	TC_3	Subtest Totals
Characteristics of Theories	3 items	3 items	2 items	8 items
Ontological Implications of Theories	3 items	3 items	2 items	8 items
Generation of Theories	3 items	3 items	2 items	8 items
Testing of Theories	3 items	3 items	2 items	8 items
Theory Choice	3 items	3 items	2 items	8 items

Legend: For Form A: TC_1 = Bohr's theory of the atom
TC_2 = Darwin's theory of evolution
TC_3 = General items

For Form B: TC_1 = Theory of plate tectonics
TC_2 = Oparin's theory of abiogenesis
TC_3 = General items

Item-Selection Results

Item-total correlation coefficients for forms A and B are listed in Tables 9 and 10. The starred items in the tables are the items with the lowest correlation coefficients for each cell (subscale x TC_n).

Starred items were deleted from the subtests and alpha coefficients were computed. Alpha values for each complete subtest and subtests with deleted items are found in Table 11.

Table 9: Product-Moment Correlation Coefficients (Item x Subtest) for Pilot Form A

	ONT	CHAR	GEN	TES	CHOICE
TC_1	.40760	-.19270*	-.10428*	-.03541	-.01015*
	.27216*	-.18162	.36333	-.15278*	.50604
	.46802	.09730	.14929	.03129	.29432
TC_2	.47248	.40897	.05160*	.02372*	.12417
	.17332*	-.04006*	.25983	.17175	.19347
	.37370	.08164	.15250	.15503	.04343*
TC_3	.57140	.09985	.19023	-.01847*	.11844*
	.34772*	-.07125*	.06281*	.15739	.35705

Legend: ONT = Ontological Implications subtest
CHAR = Theory Characteristics subtest
GEN = Theory Generation subtest
TES = Theory Testing subtest
CHOICE = Theory Choice subtest

TC_1 = Bohr's theory context
TC_2 = Darwin's theory context
TC_3 = General context

Table 10: Product-Moment Correlation Coefficients (Item x Subtest) for Pilot Form B

	ONT	CHAR	GEN	TES	CHOICE
TC_1	.24645	-.22820*	.36226	.45710	-.39145*
	.09766*	.02808	.25553*	.15922*	.18015
	.29479	-.19965	.42190	.30432	.03869
TC_2	.46967	.15920	.26631	.19323	-.29399*
	.23866*	.14713	.39901	.34901	-.29231
	.60528	-.01547*	.12187*	.06627*	.17411
TC_3	.39275*	-.05103	.16694*	.31891	.28551
	.51999	-.11707*	.27992	.17444*	-.19110*

Legend: ONT = Ontological Implications subtest

CHAR = Theory Characteristics subtest

GEN = Theory Generation subtest

TES = Theory Testing subtest

CHOICE = Theory Choice subtest

TC_1 = Plate tectonics context

TC_2 = Abiogenesis context

TC_3 = General context

Table 11: Values of Cronbach Alpha for Subtests From Pilot Forms A and B of the COST

Subtest	Form A	Form A (Deleted Items)	Form B	Form B (Deleted Items)
Theory Characteristics	.00	.34	-.09	-.08
Ontological Implications	.68	.69	.66	.68
Theory Testing	.10	.16	.53	.59
Theory Generation	.34	.49	.56	.56
Theory Choice	.45	.48	-.38	.36

The theory characteristics subtest appears the least homogeneous of any of the subtests. Three of the four computed alpha values are equal to zero or less than zero. The lack of subtest homogeneity implied by these low values of alpha would make the interpretation of scores from the subtest difficult. For this reason the theory characteristics subtest was deleted from the final form of the COST.

Alpha values for other subtests were increased by deleting items with low (item x subtest) correlation coefficients. Larger values of alpha indicate an increase in the common-factor concentration of each subtest. Improvement in the common-factor concentration of the subtests should facilitate lucid interpretation of subtest scores.

Conclusion

The final form of the COST consists of 50 items organized into the following groups: ontological implications of theories subtest

(10 items), testing of theories subtest (10 items), generation of theories subtest (10 items), theory choice subtest (10 items), and student information items (10 items). The COST and its scoring key are found in Appendix C.

The next chapter of the dissertation is concerned with the performance characteristics of the COST. This includes discussions of the validity and reliability of the instrument.

CHAPTER V

PERFORMANCE CHARACTERISTICS OF THE CONCEPTIONS OF SCIENTIFIC THEORIES TEST

Introduction

The COST is intended to measure teachers' understanding of the tentative and revisionary conception of science in terms of specific aspects of scientific theories. A question of utmost importance in the development and eventual application of the COST is whether it measures the traits it was intended to measure; that is, does it actually measure understanding of particular aspects of scientific theories? This question is answered by testing the validity of the instrument. In the first section of this chapter, the results of investigating the validity of the COST are described.

Another characteristic of the COST that is of vital significance in its successful application is its reliability. If meaningful comparisons of the results of administering the COST are to be made, then it is important to know how accurate measurements made with this instrument are. This concern is addressed by determining the reliability characteristics of the COST. Results of investigating the reliability of the COST are also presented in this chapter.

The reliability and validity of an instrument depend on characteristics of the sample to which the instrument was applied. It is appropriate in this chapter, therefore, to provide additional information on the sample of students who took the COST. This information, considered as general characteristics of the COST, includes mean time required to take the instrument and information from the student information items.

Validity Characteristics

Two approaches were used in investigating the construct validity of the COST: discrimination between contrasting groups and the multi-trait and multi-method matrix of Campbell and Fiske. The results of applying these two approaches (which are described in detail in Chapter II) are discussed below.

Discrimination Between Contrasting Groups

Several hypotheses were generated that predicted differential performance on the COST among the three groups. Verification of these hypotheses, which were described in Chapter II, would provide support for the construct validity of the COST. They are reported below:

1. On all subscales, philosophy of science students will perform according to the tentative and revisionary conception of science more than elementary education majors.
2. On the testing of theories and generation of theories subscales, chemistry students will perform according to the tentative and invention conceptions, respectively, more than elementary education majors.

3. Philosophy of science students will perform more consistently than elementary education majors on all subscales.

4. Chemistry students will perform more consistently than elementary education majors on all subscales.

The above hypotheses are of two types: subscale differences hypotheses (#1 and #2) and consistency hypotheses (#3 and #4). They are discussed in turn.

Subscale differences hypotheses. The student's t-test was used to test the statistical significance of the group differences asserted by Hypotheses 1 and 2. Because these hypotheses assert that one group will score higher than another, it was appropriate to use a one-tailed test of significance. Alpha was set at 0.1 in order to provide a somewhat liberal test of significance. T-test results of the groups referred to in Hypothesis 1 are presented in Table 12.

All differences are significant at the indicated value of alpha. Comparison of group performance on the ontological implications of theories subtest, because of the statistical equivalence of the group variances, was based on an estimate of the pooled population variance. All other comparisons were based on separate variance estimates. The results of this test support Hypothesis 1. Philosophy of science students' subtest performances are more consistent with the alternative conceptions implied by the tentative and revisionary interpretation of science than are the subtest performances of elementary education majors.

Table 12: Hypothesis 1 Test Results

Subtest	Group	Mean	S.D.	t Value	D.F.	One-Tailed Probability (α = 0.1)
Theory Choice	1 2	2.8967 2.6540	.387 .235	3.11	42.01	.002
Generation	1 2	2.4900 2.1960	.436 .299	3.26	45.54	.001
Testing	1 2	3.0533 2.4980	.489 .303	5.61	42.54	.000
Ontological Implications	1 2	2.5767 2.440	.388 .301	1.71	.78	.046

Legend: Group 1 = Philosophy of science students (30)
Group 2 = Elementary education majors (50)

T-test results of the groups addressed in Hypothesis 2 are presented in Table 13. Comparisons were made of performance on only two subtests, testing of theories and generation of theories. As discussed in Chapter II, this was done because the characteristics of chemistry students made it likely that predictable differences would be observed in performance on these two subtests. The test reveals no significant difference in performance on the generation of theories subtest. Chemistry students' scores on the testing of theories subtest were significantly greater than elementary education majors' scores as determined by a comparison of mean scores on this subtest. Thus, chemistry students' performance was more consistent with a tentative conception of theory testing than was the performance of elementary education majors. Because no significant differences were found for chemistry students' and elementary education majors' performance on the generation of theories subtest, the results of this test provide only partial support for Hypothesis 2.

Table 13: Hypothesis 2 Test Results

Subtest	Group	Mean	S.D.	t Value	D.F.	One-Tailed Probability (α = 0.1)
Generation	1 2	2.1833 2.1960	.482 .299	-.13	42.58	.897
Testing	1 2	2.9067 2.4980	.466 .303	4.29	43.88	.000

Legend: Group 1 = Chemistry students (30)
Group 2 = Elementary education majors (50)

Consistency hypotheses. Hypotheses 3 and 4, the consistency hypotheses, assert group differences in individual consistency of performance. Consistency hypotheses are important because they assert a relationship between group characteristics and performance expected on the test. A variety of measures of consistency are explored in the following discussion.

One measure of individual performance consistency is the variance of item scores for each individual, that is, the variance in the scores (1-4) an individual obtains on the eight items on a subtest. A group value of individual consistency would then be the mean of all individual values. Mean values of item variance for individuals for the three groups are provided in Tables 14 and 15. Results of t-tests of group means (Tables 14 and 15) indicate that, for all comparisons made, the item variances for elementary education majors were less than or equal to item variances for chemistry students and philosophy of science students.

Table 14: T-Test of Mean Individual Variances for Philosophy of Science and Elementary Education Students

Subtest	Group	Mean	S.D.	t Value	D.F.	Two-Tailed Probability (α = 0.1)
Testing	1 2	.5474 .5069	.370 .350	.49	78	.626
Generation	1 2	.8070 .5489	.521 .348	2.41	44.69	.020
Ontological Implications	1 2	.6944 .5444	.404 .389	1.65	78	.104
Choice	1 2	.7567 .5691	.447 .372	2.02	78	.048

Legend: Group 1 = Philosophy of science students (30)
Group 2 = Elementary education majors (50)

The discrepancy between these results and the expectation which underlies Hypotheses 3 and 4 suggests that an examination of this measure of individual performance consistency is appropriate. The rationale for investigating possible differences in individual performance consistency hinged on the assumption that possession of a particular alternative conception would be expressed in a manner consistent with that conception. Group performance consistency would then reflect an average value of some measure of an individual's consistency in answering subtest items according to a particular alternative conception.

Table 15: T-Test of Mean Individual Variances for Chemistry and Elementary Education Students

Subtest	Group	Mean	S.D.	t Value	D.F.	Two-Tailed Probability (α = 0.1)
Testing	1 2	.7052 .5069	.410 .350	2.30	78	.024
Generation	1 2	.5759 .5489	.376 .348	.33	78	.746
Ontological Implications	1 2	.8004 .5444	.474 .389	2.62	78	.112
Choice	1 2	.7811 .5691	.433 .372	2.32	78	.024

Legend: Group 1 = Chemistry students (30)
Group 2 = Elementary education majors (50)

Item variance of individuals appears to adequately capture the idea of performance consistency presented above. However, a possible ambiguity associated with the use of this measure of consistency is that the frequency of extreme item scores (strongly agree and strongly disagree) would affect the variance. Thus, two individuals' item variances might differ due to the frequency of extreme item scores even though they were identical on the basis of total number of agree and disagree responses (assuming extreme scores were collapsed to agree or disagree responses).

The above considerations suggested that differences in the certainty of response might explain the group differences in item variance. Consequently, it was hypothesized that chemistry and philosophy of science students were more certain in their responses to COST items than were elementary education majors. Certainty of response was operationalized as the number of extreme responses used in answering subtest and total test items. Philosophy of science students and chemistry students were compared to elementary education majors, using a one-tailed t-test of significance. Results of this test are reported in Tables 16 and 17. According to these results, on all subtests philosophy of science students and chemistry students responded with greater certainty than did elementary education majors.

Table 16: Certainty of Response Comparisons for Philosophy of Science and Elementary Education Students

Subtest	Group	Mean	S.D.	t Value	D.F.	One-Tailed Probability ($\alpha = 0.1$)
Testing	1 2	3.900 1.480	2.695 1.752	4.39	43.90	.000
Generation	1 2	3.300 2.120	2.493 1.881	2.24	48.81	.015
Ontological Implications	1 2	2.633 1.660	2.385 2.282	1.82	78	.037
Choice	1 2	3.667 1.700	2.591 1.810	3.66	46.13	.001

Legend: Group 1 = Philosophy of science students (30)
Group 2 = Elementary education majors (50)

Table 17: Certainty of Response Comparisons for Chemistry and Elementary Education Students

Subtest	Group	Mean	S.D.	t Value	D.F.	One-Tailed Probability (α = 0.1)
Testing	1 2	3.800 1.480	2.605 1.752	4.33	44.93	.000
Generation	1 2	2.967 2.120	2.371 1.881	1.77	78	.041
Ontological Implications	1 2	3.167 1.660	2.365 2.282	2.82	78	.002
Choice	1 2	2.667 1.700	2.106 1.810	2.17	78	.017

Legend: Group 1 = Chemistry students (30)
Group 2 = Elementary education majors (50)

The above results support the contention that the difference in item variance between elementary education majors and the other two groups is due to the greater response certainty of philosophy of science students and chemistry students. Thus, the claim that philosophy of science students and chemistry students perform more consistently than elementary education majors requires another measure of consistency.

A consistency measure that was insensitive to the frequency of extreme responses was derived by determining the absolute value of the difference between the sum of agree responses (#1 and #2) and the sum of disagree responses (#3 and #4). Values for this consistency measure could range from 10 for a completely consistent performance to

0 for a performance consisting of equal numbers of agree and disagree responses. Recalling Hypotheses 3 and 4, philosophy of science students and chemistry students were asserted to be more consistent in their performance than elementary education majors. Thus, a one-tailed t-test was used to compare group performances on this measure. The results of this test are reported in Tables 18 and 19.

Table 18: Consistency of Response Comparisons for Philosophy of Science and Elementary Education Students

Subtest	Group	Mean	S.D.	t Value	D.F.	One-Tailed Probability (α = 0.1)
Testing	1 2	5.667 3.160	3.284 2.590	3.79	78	.000
Generation	1 2	3.867 4.880	2.726 2.256	-1.80	78	.038
Ontological Implications	1 2	3.600 3.080	1.850 2.069	1.13	78	.131
Theory Choice	1 2	3.533 2.480	2.862 2.159	1.74	48.33	.088
Total Test	1 2	4.167 3.400	1.744 1.156	2.14	44.50	.019

Legend: Group 1 = Philosophy of science students (30)
Group 2 = Elementary education majors (30)

Comparisons of philosophy of science students and elementary education majors reveal that philosophy of science students are more consistent on the theory choice and testing of theories subtests. They are also more consistent on the total test. However, on the ontological implications of theories and the generation of theories subtests they were not more consistent. In fact, on the generation of theories subtest elementary education majors appear more consistent that philosophy of science students. This implies that elementary education majors have had experiences that provide them with well-developed conceptions of theory generation that are reflected in their relatively consistent performance on this subtest.

Table 19: Consistency of Response Comparisons for Chemistry and Elementary Education Students

Subtest	Group	Mean	S.D.	t Value	D.F.	One-Tailed Probability (α = 0.1)
Testing	1 2	4.800 3.160	3.428 2.590	2.26	48.87	.014
Generation	1 2	5.067 4.880	2.959 2.256	.30	49.21	.767
Ontological Implications	1 2	3.333 3.080	2.057 2.069	.53	78	.597
Theory Choice	1 2	2.467 2.480	1.634 2.159	-.03	78	.977
Total Test	1 2	3.917 3.400	1.469 1.156	1.75	78	.043

Legend: Group 1 = Chemistry students (30)
Group 2 = Elementary education majors (50)

Comparisons of chemistry students and elementary education majors show chemistry students performing more consistently only on the testing of theories subtest and the total test. On the remaining subtests chemistry students were not more consistent in their performance than elementary education majors.

In conclusion, Hypothesis 3 is only partially supported by the results of this analysis. Philosophy of science students performed more consistently than elementary education majors on only two subtests, theory choice and testing of theories. Likewise, with respect to Hypothesis 4, chemistry students performed more consistently on only the testing of theories subtest. Thus, the results of testing the consistency hypotheses provide the most support for the testing of theories subtest. These results provide some support for the theory choice subtest and little support for the remaining subtests.

The Multi-trait and Multi-method Matrix

It was hypothesized that, if the constructs underlying the test were valid, then the test would satisfy the first three requirements for construct validity specified by Campbell and Fiske. These requirements are summarized below:

1. The validity coefficients of a test (coefficients of correlation between trait scores measured with different methods) must be significantly greater than zero.

2. The validity coefficients must be significantly greater than correlations between different trait scores using the same method.

3. The validity coefficient for a given trait must be significantly greater than the correlation between measurements of this trait and measurements of all other traits with any other method.

The requirements (as discussed in more detail in Chapter II) are appropriate to testing the validity of the COST constructs.

Requirement 1 is based on the expectation that there should be significant agreement between measurements of the same construct with different methods (convergent validity). Requirements 2 and 3 are based on the expectation that there should be little agreement between measurements of different constructs (discriminant validity).

In order to construct a multi-trait and multi-method matrix, trait-method scores for all subjects were computed. These scores resulted from summing the two item scores that shared both subtest (trait) and theoretical context (method). Pearson product-moment correlations were then determined for all pairs of trait-method scores. The resulting matrix is found in Appendix D.

The results of applying Campbell and Fiske's first requirement to the validity coefficients ($\alpha = 0.1$) are reported in Table 20.

Table 20: Ratio of Significant Validity Coefficients to Total Number of Validity Coefficients

Subtest	Ratio
Ontological Implications	6/10
Testing of Theories	10/10
Generation of Theories	8/10
Theory Choice	6/10

This requirement is satisfied most by the testing of theories subtest, followed by the generation of theories subtest. Both theory choice and the ontological implications of theories subtests satisfy this requirement least.

The results of applying Requirements 2 and 3 of Campbell and Fiske's validation procedure are presented in Table 21. The values reported are the percentages of the appropriate correlation coefficients that were significantly less than the validity coefficients. Comparison of values of the validity coefficients with other correlation coefficients in the matrix required that the correlation coefficients be transformed into Fisher's Z values and confidence intervals around these values be established.

Table 21: Proportion of Significant Matrix Correlation Coefficients

Subtest	Smaller Multi-trait and Monomethod Coefficients[a] (%)	Smaller Multi-trait and Multi-method Coefficients[a] (%)
Ontological Implications	47	43
Testing of Theories	48	47
Generation of Theories	50	53
Theory Choice	20	29

[a]Significantly smaller than validity coefficient at 0.1 level.

All subtests except theory choice have approximately the same proportion of multi-trait/monomethod and multi-trait/multi-method coefficients that are significantly smaller than the relevant validity coefficients. The percentage of the relevant coefficients for the theory choice subtest seems comparatively low and, therefore, anomalous.

Reflection on the results of applying Campbell and Fiske's requirements for construct validation suggests that these results provide relatively poor support for the ontological implications of theories and theory choice subtests. Both of these subtests have the least number of significant validity coefficients (Table 20). And the theory choice subtest satisfies least Criteria 2 and 3 of Campbell and Fiske. An examination of the detailed structure of these subtests may prove useful in understanding the results of the validity tests.

Each item in the COST was based on a statement from the philosophic framework presented in Tables 1 through 5 (Chapter II). For example, item 5 is based on a statement from the conclusive view of theory testing. For all subtests, the statement numbers and the number of items based on each statement are presented in Table 22. For example, one item in the theory testing subtest is based on the first statement of the tentative interpretation of this subtest. This is represented as T1(1). Statements that are paired opposites (e.g., C3 and R3) are considered as one statement. Also, some items involve more than one statement (e.g., R1,2 represents an item based on statements #1 and #2).

On the basis of the number of statements used in a subtest, the theory choice subtest appears the most heterogeneous. Even though the statements in the theory choice subtest are related to the same aspect of scientific theories, it is conceivable that its heterogeneity accounts for this subtest's relatively poor performance on Campbell and Fiske's validity tests. However, this explanation is inadequate to account for the low proportion of significant validity coefficients for the ontological implications of theories subtest (Table 20). Both this subtest and the generation of theories subtest are based on the same number of statements. And yet, the generation of theories subtest is superior in its proportion of significant validity and coefficients.

Table 22: Statement Numbers and the Number of Items Based on Each Statement

Subtest	Statement Numbers and (Items per Statement)	Total Statements
Theory Choice	R1,2(1) R4(1) R5(1) C3(1) C1(2) R2(1) R2,5(1) C1,2(1) R3(1)	5
Ontological Implications	I1,3(1) I4(2) R1(4) I2(1) R1,3(2)	4
Testing	C2(4) C1(5) T1(1)	2
Generation	Id1,2(2) Id3(3) In2(1) Id4(1) In1,2(1) In3(2)	4

Legend: R = Revisionary, C = Conclusive
C = Cumulative, T = Tentative
I = Instrumentalism, Id = Induction
R = Realism, In = Invention

A possible explanation of this discrepancy is that, in spite of the use of the same number of statements in the ontological implications of theories and generation of theories subtests, these subtests still differ in their semantic heterogeneity. The considerable variation in the complexity of the statements is consistent with this explanation. For example, consider this generation of theories statement (In3):

> How a theory is generated is irrelevant to its usefulness.

The simple structure of this sentence contrasts with the complexity of the following statement from the ontological implications of theories subtest (I4):

> Because different and incompatible theories may be used by scientists in their search, it is not appropriate to claim that any of them describe what is true.

This digression into the structure of some of the COST subtests has served to illuminate some of the characteristics of the subtests that influence their performance. Differences exist among the subtests that are reflected in their performance on the validity tests of Campbell and Fiske. These differences appear to result from different amounts of subtest hererogeneity. This understanding of subtest characteristics should prove useful in interpreting performance on the COST. For example, an item analysis of performance on the theory choice subtest might reveal consistent patterns of response to particular statements in that subtest that would not be reflected in the subscale score. Such an analysis would be appropriate because of the demonstrated heterogeneity of the subtest and would provide information useful in characterizing a teacher's understanding of the issues embodied in that subtest.

In conclusion, results of applying the validity tests of Campbell and Fiske provide varying amounts of support for the construct validity of the COST subtests. The testing of theories and generation of theories subtests have the strongest support on the basis of the proportion of validity coefficients that were significantly greater than zero (Table 20). These subtests demonstrate considerable convergent validity. All subtests, with the exception of theory choice, demonstrate approximately equivalent amounts of discriminant validity as reflected in the percentages of matrix correlation coefficients smaller than the validity coefficients (Table 21).

Validity Conclusions

The results of applying two procedures to the investigation of the construct validity of the COST are summarized in Table 23. Strongest support exists for the testing of theories subtest, even though some support exists for all subtests. The significance of this validity support is that it provides justification for the claim that the COST measures understanding of particular aspects of scientific theories. Because of the relevance of the validity support in providing justification for the intended use of the COST, additional comment on this support is necessary.

According to the validity criteria of this study, the testing of theories subtest has the strongest support. All of the group difference hypotheses for this subtest were confirmed. The convergent validity requirement of Campbell and Fiske was completely satisfied. The only weakness in the validity support for this subtest was incomplete satisfaction of the requirement of discriminant validity. This weakness is not considered serious for two reasons. First, some relationship should exist between the COST aspects because they are all aspects of scientific theories and, therefore, related. And second, because no generally agreed upon criteria are available for evaluating the sufficiency of discriminant validation evidence, a relative judgment is required. Relative to the other COST subtests, the testing of theories subtest is approximately equivalent to the best three subtests in satisfying the requirement of discriminant validity.

Table 23: Summary of Construct Validity Evidence

Subtest	Group Difference Hypotheses				Rank Order of Satisfaction of Requirements of Campbell and Fiske
	1	2	3	4	
Testing of Theories	+	+	+	+	1
Generation of Theories	+	-	-	-	2
Ontological Implications of Theories	+	NT	-	-	3
Theory Choice	+	NT	+	-	4

Legend: + = Hypothesis supported by evidence
- = Hypothesis not supported by evidence
NT = Hypothesis not tested

At the opposite extreme in the extent of its validity support is the theory choice subtest. This subtest demonstrates the least discriminant validity, and, along with the ontological implications of theories subtest, this subtest satisfies the requirement of convergent validity least. And yet, the theory choice subtest has some validity support. Sixty percent of its validity coefficients are significantly greater than zero. And two group difference hypotheses concerning this subtest were confirmed. As discussed earlier, the conceptual heterogeneity of the theory choice subtest may explain its relatively weak validity support. At the same time, a review of instruments that address the domain of the nature of science reveals that there are no instruments that assess understanding of the issues embodied in the theory choice subtest. Consequently, with the caveat that performance on this subtest should be interpreted with an appropriate awareness of its heterogeneity, the writer claims that the importance of assessing teachers' understanding of the tentative and revisionary conception justifies use of the theory choice subtest.

The two remaining subtests, generation of theories and ontological implications of theories, have validity support intermediate in strength between the two extremes just described. The generation of theories subtest has relatively strong convergent validity support (80% of testing of theories). The ontological implications of theories subtest has better discriminant validity support than the theory choice subtest. In the absence of accepted criteria for evaluating the adequacy of the validity support for these subtests, the writer claims that there is sufficient evidence to justify use of the generation of theories and ontological implications of theories subtests.

Reliability Results

Cronbach alpha reliability coefficients and standard errors of measurement were computed for all subtests. Values for these variables are reported in Table 24. The range of values for alpha for elementary education majors are lower than values for the other two groups. This can be explained by the greater homogeneity of the group of elementary education majors (as evidenced by the relatively low values of the standard deviation).

The magnitude of the standard error of measurement does not depend on the sample's homogeneity. "The measurement of a variable for a single individual takes place with a certainty which is independent of the homogeneity of the sample in which he is included" (Magnusson, 1967, p. 82). This suggests that in administering the COST the standard error of measurement may be a better predictor of the instrument's reliability than obtained values of the reliability coefficient. The relatively low values of the standard error of measurement (all less than 0.3 for a scale interval of 3) indicate an adequate amount of certainty may be attributed to COST scores. This claim is made relative to the intended use of the COST in inferring conceptions of aspects of scientific theories. A standard error of this magnitude would not result in a confidence interval around an obtained subscore that would include both extremes of the four-point scale. Thus, associating a particular conception with a score would not have to be changed due to size of the confidence interval around that score. More will be said about this in the following chapter.

Table 24: COST Reliability Results

Group	Total Test			Ontological Implications of Theories			Testing of Theories			Generation of Theories			Theory Choice		
	α	S	SE	α	S	SE	α	S	SE	α	S	SE	α	S	SE
Elementary Education Majors (50)	.493	.157	.112	.482	.301	.217	.503	.303	.253	.437	.299	.224	.126	.235	.220
Philosophy of Science Students (30)	.821	.302	.128	.638	.388	.233	.796	.489	.221	.604	.436	.274	.575	.387	.252
Chemistry Students (30)	.728	.236	.123	.679	.411	.233	.774	.466	.222	.779	.482	.226	.350	.287	.231

Legend: α = Cronbach alpha reliability coefficient
S = Sample standard deviation
SE = Standard error of measurement

General Characteristics of the COST

Administration Time

An estimate of the time required to complete the test was obtained by asking the sample of elementary education majors to record the time spent on the test on their answer sheets. The mean time required by 32 members of the sample who supplied this information was 28.3 minutes. This time requirement would permit administration of the COST during a normal 50-minute class period.

Student Information Items

The reliability and validity of the COST depend on the characteristics of the sample to which the instrument was administered. Some of these characteristics were described previously in Chapter II. Additional information was obtained from the student information items of the COST.

One type of student information items asked student to assess their subject-matter knowledge. The results of administering these items are included in Table 25. Both philosophy of science students and elementary education majors indicated that their knowledge of Darwin's theory of evolution was greater than for the other topics. Not surprisingly, chemistry students rated Bohr's theory of the atom as the topic about which they had the greatest knowledge. Knowledge of geological theories was rated lowest by chemistry students and philosophy of science students. Elementary education majors rated their knowledge of all topics except Darwin's theory as "somewhat competent" to "slightly competent."

Table 25: Self-Assessment of Subject-Matter Knowledge

Knowledge Assessed	Philosophy of Science Students	Chemistry Students	Elementary Education Majors
Bohr's Theory	3.4 (.89)	2.3 (.75)	3.8 (1.2)
Darwin's Theory	2.4 (.86)	2.8 (.87)	2.5 (.93)
Geological Theories	3.9 (1.1)	4.0 (1.0)	3.6 (1.1)
Theory of Abiogenesis	3.5 (.82)	3.4 (.97)	3.9 (1.4)
Philosophy of Science	3.1 (.68)	3.2 (.97)	3.3 (1.4)

Mean Values (Standard Deviation)

Scale: Mastery 1
Highly Competent 2
Somewhat Competent 3
Slightly Competent 4
No Knowledge 5

The second set of student information items assessed personal bias, which was defined as influences, other than subject-matter knowledge, that may affect performance on the COST. Results of administering these items are found in Table 26. Both philosophy of science students and elementary education majors expressed "strong" to "moderate" bias toward Darwin's theory of evolution. Chemistry students' ratings of bias for all topics were "moderate" to "weak."

Personal bias and subject-matter knowledge may be used to investigate factors that influence performance on the COST. Since this is not the intention of this study and since no single procedure is readily applicable, procedures for pursuing this investigation are not addressed.

Table 26: Self-Assessment of Personal Bias

Personal Bias	Philosophy of Science Students	Chemistry Students	Elementary Education Majors
Bohr's Theory	3.0 (1.2)	3.4 (.97)	3.4 (1.4)
Darwin's Theory	2.5 (1.1)	3.5 (1.0)	2.6 (1.2)
Geological Theories	3.3 (1.4)	3.9 (.94)	3.1 (1.3)
Theory of Abiogenesis	3.0 (1.4)	3.7 (.96)	3.5 (1.5)

Mean Values (Standard Deviation)

Scale:		
	Complete	1
	Strong	2
	Moderate	3
	Weak	4
	None	5

Concluding Comments

The validity and reliability of the COST, which have been discussed in this chapter, are sufficient for the application of this instrument in its intended domains. The COST was developed to measure teachers' understanding of particular aspects of scientific theories that relate to the tentative and revisionary conception of the nature of science. The dependability characteristics of the COST justify the claim that the COST measures the above attributes and that a reasonable amount of certainty may be attributed to these measurements. In the succeeding chapter, the results of applying the COST to the groups addressed in this study are described.

CHAPTER VI

INFERRING CONCEPTIONS OF SCIENTIFIC THEORIES BY USING THE CONCEPTIONS OF SCIENTIFIC THEORIES TEST

Introduction

The intention of this work was to develop a means of inferring teachers' conceptions of particular aspects of the nature of science. Previous chapters have presented discussions of the basis and justification of the instrument, instrument-development procedures, and instrument characteristics. This chapter of the dissertation focuses on the results of applying the COST to three different groups: elementary education majors, chemistry students, and philosophy of science students. Inferences concerning conceptions of scientific theories held by members of these groups are discussed.

Conceptions of Scientific Theories Held by Elementary Education Students

The primary focus of this discussion is on the conceptions held by elementary education students. This is because the intention of this work was to develop a test that could be used to infer teachers' conceptions of particular aspects of the nature of science. Secondarily, and as a means of augmenting the description of elementary education students' conceptions, the conceptions of chemistry and philosophy of science students are discussed.

Frequencies of Recoded Subtest Performance Scores

For this analysis, the subtest scores were recoded in order to facilitate pattern recognition. It was assumed that subtest scores that ranged from 1.0 < 2.3 indicated performance to some degree consistent with the alternative subtest conception associated with one side of the scale. Scores in this range were recoded "1." Scores in the 2.3-2.7 range were recoded "2" and considered indicative of no particular alternative conceptions (or a conception not assessed by the COST) and, therefore, an indeterminate conception. And scores in the 4.0 > 2.7 range were recoded "3," as an indication of performances consistent with the alternative subtest conception associated with the opposite side of the scale (see Figure 7). The score boundaries that result in recoded scores have practical significance. The magnitude of the standard errors of measurement for the subtests suggests that an observed score reflecting a particular alternative conception would not result for a true score due to the opposite alternative conception. For example, an observed score of 2.8 on the theory testing subtest would be due with 90% certainty to a true score found in the range 2.4-3.2.

The subtest score range that results in an indeterminate designation (a recoded score of 2) deserves additional comment. This range, 2.3 to 2.7, occurs for two different patterns of subtest items scores. One pattern consists of equal numbers of ones and fours or twos and threes. This pattern obviously reflects a response pattern inconsistent with any particular subtest alternative. The second pattern of

Subtests Score Range	Recoded Score	Ontological Implications	Testing of Theories	Generation of Theories	Theory Choice
2.3	1	Realist	Conclusive	Induction	Objectivist
2.3-2.7	2	Indeterminate	Indeterminate	Indeterminate	Indeterminate
2.7	3	Instrumentalist	Tentative	Invention	Subjectivist

Figure 7. Subtest conceptions and recoded subscores.

Subtests Score Range	Recoded Score	Ontological Implications	Testing of Theories	Generation of Theories	Theory Choice
2.3	1	Realist	Conclusive	Induction	Objectivist
2.3-2.7	2	Indeterminate	Indeterminate	Indeterminate	Indeterminate
2.7	3	Instrumentalist	Tentative	Invention	Subjectivist

Figure 7. Subtest conceptions and recoded subscores.

responses that produces scores within the indeterminate range consists of several intermediate item scores (2 or 3) plus a few extreme scores (1 or 4). For example, an item response pattern such as 2,2,2,2,2,2, 2,2,3, and 4 would produce a score of 2.3. Such a score would (according to the assigned boundary values) indicate an indeterminate performance. And yet, only two items' scores are inconsistent with the remainder. This example illustrates that the chosen boundary values result in indeterminate designations for item response patterns that are minimally inconsistent. That is, this represents a rather stringent criterion. With these comments concerning the significance of the recoded subtest scores in mind, the analysis of COST performance based on these scores may be discussed.

Conceptions

Frequencies of recoded values of subtest performance are presented in Table 27. Inferences concerning elementary education students' conceptions of subtest aspects were based on these data.

On the generation of theories subtest, 78% of this sample of elementary education students performed in a way consistent with an induction conception of theory generation (refer to Figure 7). The induction conception is a naive view of how theories are generated (Hempel, 1966, p. 11). In contrast, only 37% of the philosophy of science students, who because of their experience with the philosophy of science would be expected to have a more sophisticated understanding of theory generation, perform in a way consistent with the induction conception.

Table 27: Frequencies of Recoded Values of Subtest Performance

Subtest		Elementary Education Majors	Philosophy of Science Students	Chemistry Students
Generation of Theories	1	78%	37%	63%
	2	17%	23%	17%
	3	6%	40%	20%
Testing of Theories	1	28%	3%	7%
	2	44%	20%	30%
	3	28%	77%	63%
Ontological Implications of Theories	1	36%	33%	40%
	2	44%	27%	20%
	3	20%	40%	40%
Theory Choice	1	10%	7%	30%
	2	44%	20%	30%
	3	46%	73%	40%

Performance on the testing of theories subtest indicates no strong preference by elementary education students for any particular conception. A plurality of this group performed indeterminately. This performance contrasts with the performance of chemistry students and philosophy of science students. Majorities of both these groups performed according to the tentative conception of theory testing.

Preservice elementary teachers' conceptions of the ontological implications of theories may be inferred from their scores on this subtest. Forty-four percent of this group performed in an indeterminate way. Thirty-six percent performed according to a realist conception of this subtest. Only a fifth of this group indicated an understanding of the instrumentalist conception of the ontological implications of theories. This performance contrasts with the scores of both chemistry and philosophy of science students (40% instrumentalist).

One of the more surprising results of this analysis is the relatively large proportion (46%) of the elementary education students whose performance is consistent with the revisionary conception of theory choice. This controversial conception of theory choice, although widely discussed in philosophic circles, is not a view that is commonly addressed by science educators. The proportion of elementary education students performing in accordance with this view is greater than the corresponding proportion of chemistry students (40%). And, once again, probably because of course experience, the highest proportion of students who performed consistent with the revisionary view is from the philosophy of science students (73%).

Inferring a Tentative and Revisionary Conception of the Nature of Science

As discussed in detail in Chapter II, a major goal motivating the work described in this dissertation was development of an instrument that could be used to infer teachers' understanding of the tentative and revisionary conception of the nature of science. This conception of science emphasizes the tentativeness of scientific knowledge and the revision of that knowledge in response to changing theoretical contexts. The relationship between this conception of science and the subtest alternatives was described in Chapter II. The following alternative subtest (Figure 8) conceptions are consistent with the tentative and revisionary conceptions of the nature of science.

Subtest	Alternative Conception
Testing of Theories	Tentative
Generation of Theories	Invention
Ontological Implications of Theories	Instrumentalist
Theory Choice	Subjective

Figure 8. Alternatives consistent with the tentative and revisionary conception.

The problem of inferring understanding of the tentative and revisionary conception may be approached in a variety of ways. One approach is to determine the percentages of each group that had recoded scores of three on all subtests. These percentages are listed in Table 28. According to the criterion of scores of three on all subtests, very small proportions of all groups demonstrated evidence of possession of the tentative and revisionary conception of the nature of science.

Table 28: Subtest Scores and Total Scores as Evidence of Understanding the Tentative and Revisionary Conception

Groups	(%) All Subscores = 3	(%) Total Score > 2.7
Elementary Education Students	0	4
Philosophy of Science Students	13	43
Chemistry Students	3	27

The small percentage of philosophy of science students (13%) who performed in accordance with this conception is surprising. All of these students have read The Structure of Scientific Revolutions, which contains discussions of many of the ideas central to a tentative and revisionary conception of the nature of science. Although it is conceivable that students understood this conception but did not accept it, I also think it likely that this criterion for demonstrating possession of a tentative and revisionary conception is too stringent. This is because the relationship between this conception and particular subtest alternatives is not exact. The tentative alternative of the testing of theories subtest relates most directly to the tentative and revisionary conception. Other subtest alternatives relate less precisely to this conception. This makes the requirement that evidence of a tentative and revisionary conception consist of recoded scores of three inappropriately strict.

Consequently, a different criterion on which to base distinctions among COST performances was also used. This criterion required that the recoded total score, a measure of average performance across all subtests, be greater than 2.7. The value, 2.7, was used to distinguish recoded subscores of 2 from 3, and represented in this application, the boundary between COST performances consistent with a tentative and revisionary conception of science and those that were not. The results of applying this criterion to performance results are in Table 28. These results agree with the expectation that a sizable percentage of the sample of philosophy of science students possess a tentative and revisionary conception of the nature of science. For this reason, and because the total score is a measure of overall performance less dependent on a minimum level of performance on each subtest, the total score is the preferred basis for inferring that a tentative and revisionary conception accounts for COST performance.

Summary and Conclusion

In this chapter, inferences based on the results of applying the COST to three different groups were discussed. The following findings concerning conceptions of scientific theories held by a sample of elementary education students are obtained:

1. A majority have an induction conception of theory generation.
2. A plurality have an indeterminate conception of theory testing.
3. A plurality have an indeterminate conception of the ontological implications of theories. Over a third have a realist conception of this aspect of scientific theories.
4. A plurality have a subjective conception of theory choice.
5. A very small proportion (4%) have conceptions consistent with a tentative and revisionary interpretation of science.

Before discussing possible explanations for the possession of these conceptions by elementary education students, it is necessary to address the generality of the conclusions concerning these conceptions. What justification exists for asserting that the conclusions of this study apply to other samples of the population of elementary education students?

The sample of elementary education students consisted of all students enrolled in a course in physical science at Michigan State University. This course is required of all elementary education majors. Thus, this sample is representative of the population of elementary education majors at Michigan State University. The test administration conditions were traditional. Interaction of these conditions with testing would be negligible. Consequently, generalization of this study's conclusions to the population of elementary education students at Michigan State University is justified. Generalization to populations of preservice elementary teachers in other settings would require judgments concerning the similarity of those populations to the population addressed in this study.

This chapter concludes with a discussion of some possible explanations for the conceptions of scientific theories held by elementary education students. These students typically have had little experience with science. Because of this lack of experience they tend to have a naive, simplistic view of science. The induction view of theory generation held by this group is consistent with a naive understanding of the scientific enterprise. Also, the indeterminate performance of elementary education students on the testing of theories and ontological implications of theories subtests suggests that, because of their lack of familiarity with science, they are confused by the issues represented in these two subtests.

Another explanation of elementary education student performance on COST, consistent with an explanation based on their lack of science experience, is that students see science as a "rhetoric of conclusions"

(Schwab, 1962, p. 24). In other words, these students conceive of science as an accumulation of facts, with little awareness of their coherence and organization, and little awareness of the inquiry that gave rise to those facts. This view of science tends to neglect the conceptual and constructionist nature of scientific knowledge. Consequently, all knowledge is reified and seen as a collection of immutable facts.

Students who view science as a "rhetoric of conclusions" should have conceptions of COST subtests that neglect its conceptual and changing nature. The alternative conceptions in Table 29 are consistent with this view.

Table 29: Subtest Alternatives Consistent With a "Rhetoric of Conclusions" Interpretation of Science

Subtest	Alternative Conception	Elementary Education Students (%)
Generation of Theories	Induction	78
Testing of Theories	Conclusive	28
Ontological Implications of Theories	Realist	36
Theory Choice	Objective	10

Elementary education students' performance on COST is consistent with a "rhetoric of conclusions" view of science on some subtests. They have an induction conception of theory generation. In addition, even though a plurality of the sample were indeterminate in their conception of the ontological implications of theories, over a third expressed a realist conception of this subtest.

Performance on the theory choice subtest is inconsistent with a "rhetoric of conclusions" view of science. Even though a majority of elementary education majors were either "cumulative" or indeterminate in their understanding of theory choice (Table 27), the large proportion (46%) answering according to a "revisionary" conception is not easily explained.

The explanations of COST performance offered in the preceding discussion are admittedly speculative and incomplete. They were offered as possibilities that might suggest research projects which would focus on explanations of COST performance. This, and other research problems that may be addressed using COST, will be discussed in the following, concluding chapter of the dissertation.

CHAPTER VII

SUMMARY AND DISCUSSION

Summary

The Conceptions of Scientific Theories Test

Test development. One major purpose of this study was to construct a reliable and valid test for elementary and secondary school teachers of science that would assess their conceptions of some philosophic aspects of scientific theories. More specifically, the COST was designed to satisfy the following criteria: (1) the COST is sensitive to two alternative conceptions of selected philosophic aspects of scientific theories, and (2) the COST may be used to infer possession of a tentative and revisionary conception of the nature of science.

A test-construction framework was developed that consisted of five philosophic aspects of scientific theories (i.e., testing, generation, characteristics, ontological implications, and choice). Two alternative conceptions of each aspect were described, and items were written to discriminate between these alternative conceptions. Some items were adapted to the contexts of particular scientific theories by prefacing them with a brief description of a scientific theory and episodes drawn from its history. This resulted in an equal distribution of items between the following five groupings: Bohr's theory of the atom, Darwin's theory of evolution, Oparin's theory of abiogenesis, the theory of plate tectonics, and nontheoretical items.

Eighty items were written and divided equally between two pilot forms (A and B) of the test. These pilot forms were administered to 56 college physical science students during the summer of 1978. These students were primarily elementary education majors. Twenty-nine students took form A, and 27 took form B. The results of the pilot administration were used to select the items that related most strongly to the COST subtests which were based on the aspects of theories in the test-construction framework.

The final form of the instrument, which contained 50 items, was administered to three groups: 50 elementary education students, 30 chemistry students, and 30 philosophy of science students. Data collected from administration of the COST to these groups were used to determine the performance characteristics of the instrument.

<u>Dependability characteristics</u>. The construct validity of the COST was investigated using two approaches: discrimination between contrasting groups and the multi-trait and multi-method matrix of Campbell and Fiske. Subtest construct validity was supported by the results of these investigations. The relative strength of the validity evidence (obtained from both approaches) is as follows: testing of theories > generation of theories > theory choice $\simeq$ ontological implications of theories.

Cronbach alpha reliability coefficients and standard errors of measurement were computed for the test and all subtests. Even though a considerable range of alpha was obtained (.126 to .796), the range of values of the standard error (.112 to .274) indicates that an adequate degree of accuracy may be attributed to test scores.

Conceptions of scientific theories. The second purpose of this study was to use the COST to determine preservice elementary science teachers' conceptions of science. Subtest performances were analyzed to determine conceptions of the aspects of scientific theories assessed by the COST. And total test performance was used as a measure of preservice teachers' possession of a tentative and revisionary conception of the nature of science.

Only 4% of the sample of elementary education students tested performed in a way consistent with the tentative and revisionary conception of science. Their performance, in general, reflected a naive view of science. A majority expressed an induction conception of theory generation. On the remaining subtests, 44% of the sample performed inconsistently, indicating confusion concerning the subtest alternatives.

In contrast, no more than 27% of the sample of philosophy of science students performed inconsistently on any subtest. Similarly, for chemistry students, no more than 30% of the sample performed inconsistently. Forty-three percent and 27% of the philosophy of science and chemistry students, respectively, performed according to a tentative and revisionary conception of the nature of science.

Discussion

Contributions of the Study to Educational Research and Practice

Reflection on the results of this study suggests two major contributions to education: one, provision of an instrument for assessing teachers' conceptions of particular aspects of the nature of science; and two, description of some characteristics of preservice elementary science teachers that indicate deficiencies in their education.

Provision of a test for assessing teachers' conceptions of scientific theories is a significant contribution to research on science teacher education. The NARST-NIE Commission on Research in Science Education has recommended the development of reliable instruments for "assessing the conceptions and skills of teachers regarding science" (Yager, 1978, p. 105). This recommendation is based on the recognition that teacher characteristics influence both classroom teaching and interactions.

Administration of the COST to the sample of preservice elementary teachers described in this study revealed that very few of them possessed a tentative and revisionary conception of the nature of science. Acceptance of the importance of understanding this conception leads to the conclusion that the population of preservice elementary teachers at Michigan State University is deficient in their understanding of the nature of science. Acceptance of the assumption that the sample of preservice teachers used in this study is representative of the national population suggests that the tentative and revisionary conception is not understood by more preservice elementary teachers.

Implications for Education

The following educational implications derive from a somewhat liberal generalization of the results of this study:

1. Preservice elementary science teachers fail to understand the tentative and revisionary conception of science. Because of this deficiency, the education of these teachers should be modified to address this important goal of education.

2. The general public, as represented by the sample of non-science-oriented elementary education majors, fails to understand the tentative and revisionary conception of science. This deficiency has implications for K-12 science education. Specifically, this investigator proposes successful adoption and implementation of K-12 science curricula that emphasize understanding the tentative and revisionary conception of science.

One particular manifestation of elementary education majors' failure to understand the tentative and revisionary conception of science is their expression of an "induction" view of theory generation. This interpretation, in neglecting the role of conceptual invention, reflects a naive conceptualization of the process of theory generation. An appropriate response to the deficiency is giving more attention in instruction to the accurate description of the development of scientific knowledge. This would contrast with the traditional textbook presentation of scientific development which, in reconstructing scientific developments in terms of current conceptions, neglects the conceptual revisions which accompanied those development (Kuhn, 1970a, p. 140). Some authors have suggested the use of carefully selected original scientific papers as curriculum materials to accomplish this end (Ravetz, 1971; Schwab, 1962). In any case, in the light of this study giving more attention to the characteristics of theory generation is a desideratum of importance in elementary science teacher preparation.

Likewise development of understanding of other aspects of the tentative and revisionary conception of science could result from explicit attention to the issues raised by these aspects. Courses for elementary education majors could be improved by devoting more attention to these philosophic issues.

It is fairly obvious by now that, in spite of the implication of the tentative view of science, the writer is convinced that the tentative and revisionary conception is an important interpretation of the nature of science. This emphasis should not be construed as reflecting an attitude that this conception is the only correct view of science. Rather, this conviction derives both from an assessment of the supporting arguments and from a belief in the social and educational significance of understanding this conception.

The importance of a scientifically informed public was addressed earlier in this dissertation. This importance is due to two factors: (1) the requirement for public support of the scientific enterprise, and (2) the requirement for public participation in a society that is increasingly influenced by science and technology. Schwab (1962) has contended that these social requirements are best met by understanding science as a revisionary process whose knowledge claims are tentative. And yet, this study has revealed that the general public, as represented by the sample of elementary education majors, fails to understand this important conception of the nature of science. One implication of this deficiency is pursued in the following discussion.

Schwab has cogently argued that, for the public to understand science as it is, that is, a science whose knowledge is tentative and revisionary, then "science as inquiry" must be emphasized. From the proliferation of science curricula that emphasize, at least nominally, science as inquiry, one could conlude that Schwab's advice has been heeded. And yet, assuming the best for these curricula (both in being adequate to their goals and being implemented successfully), adoption has been less than adequate. The 1977 National Survey of Science, Mathematics, and Social Studies Education (Weiss, 1978, p. 78) reported the following figures for use of federally funded curriculum materials:[1]

1. Thirty-one percent of the national sample of school districts are using one or more of the K-6 science curricula.[2]

---

[1]Starting in 1956 the National Science Foundation has funded over 30 science curriculum-development projects. Most of the inquiry-oriented curricula are due to this support.

[2]These percentages must be interpreted with reservation. A district was considered to be using a federally funded curriculum if no more than <u>one</u> classroom in the district used these materials.

2. Sixty percent of the national sample of school districts are using one or more of the 7-12 science curricula.

The adoption figures for elementary curricula are discouraging. And, in spite of the more promising figures for grades 7 to 12, there is little room for optimism at that level. The "Case Studies in Science Education" prepared for the National Science Foundation reported the following:

> One of the more important findings of this case study project was that, despite considerable contact with legacies of the NSF-sponsored curriculum projects and with inservice programs dedicated to the promotion of student inquiry, very little inquiry teaching was occurring in science, math, and social science in the eleven sites. Lessons typically were organized by teachers around printed or dittoed materials. Problems were worked by the students, following the example set by the teacher, who helped out when an obstacle was met, but who gave little encouragement to go beyond the problem or to question an implication (Stake & Easely, 1978, pp. 12-14).

Even though these case studies were done in only 11 sites, these sites were chosen to provide a representative sample of the school districts from throughout the continental United States. Thus, findings reported in these studies are credible testimonies to the absence of teaching "science as inquiry" in American schools.

One part of the problem of educating the public to understand the tentative and revisionary conception of science then becomes how to successfully implement and adapt inquiry-oriented curricula to the schools. The importance of adaptation is critical. Shulman and Tamir (1977) have stated:

> Moreover, our experience with studies of implications of these curricula has made it clear that the key to successful curriculum development and application in the schools is an understanding of how one adapts national or broad-scale curricula to local conditions. Adaptation occurs whether or not it is planned, so it had better be anticipated (p. 10).

If the science curriculum is to achieve its intention of communicating a particular view of science, then it is obligatory that the intention not be contravened by its adaptation and use in the schools. Studies of science curriculum adaptation and use would provide valuable information to use in increasing the successful adoption and implementation of science curricula that emphasize understanding the tentative and revisionary conception of science.

Limitations of This Investigation

A significant limitation of this work is the ambiguity associated with interpretations of subtest and total scores in the intermediate range. Scores are assigned to items, subtests, and total test with the assumption that a particular conception may be associated with that score. This assumption presents no difficulty at the level of the item. Responses to items may be interpreted dichotomously so that any item response is associated with one of two alternative conceptions. However, at the level of the subtest and total test, ambiguity exists when considering scores in the intermediate range. Scores that range from 2.3 to 2.7, which for the frequency analysis of subtest performance were recoded 2, imply a performance inconsistent with either of the alternative conceptions represented by the subtest. This inconsistency, however, appears amenable to at least two interpretations: (1) that it is indicative of a confused understanding of the issue assessed by the subtest items; and (2) that it is evidence of a conception not represented by the subtest alternatives.

Both of these interpretations, which contribute to the ambiguity associated with interpretations of scores in the intermediate range, are amenable to investigation. Discussion of the research possibilities implied by these interpretations will be deferred until the section on suggestions for future research.

A second limitation of this study is the heterogeneity of the samples that were the subjects of the test administration. Heterogeneity was especially apparent in the sample of philosophy of science students. Members of this sample were from two different philosophy of science classes and selected at three different times during the instructional term. The significance of this heterogeneity lies in attempting to explain test performance. Because of the variability in experience of the members of the philosophy of science students sample, it is difficult to make specific claims about the relationship between their experience and their test performance. Use of a more homogeneous sample would not only obviate this difficulty, but would, in all likelihood, provide stronger evidence in support of the construct validity of the subtests.

Suggestions for Future Research

The research possibilities suggested by this study are of two types: first, research intended to improve the COST; and second, research utilizing the COST in addressing problems in science education.

The first research possibility comprises two concerns: the validity of the theory choice subtest and ambiguities in interpreting

COST performance. It was suggested earlier that the relatively poor construct validity support for the theory choice subtest was due to the heterogeneity of this subtest. In the light of this explanation, efforts to improve the construct validity of this subtest should focus on its heterogeneity. One approach for improving the subtest's homogeneity is to reduce the number of statements on which items are based. A re-examination of the statements used in this subtest would hopefully lead to a reduction in their number while maintaining its identity. An additional reason for re-examining the statements of the theory choice subtest is that recent philosophical discussions (Laudan, 1977) suggest that the alternative interpretations of theory choice represented in COST may be simplistic. Efforts to incorporate into the theory choice subtest insights gleaned from these discussions could improve the content validity of this subtest.

A second concern in improving the COST is the ambiguity associated with intermediate-range subtest scores. This problem, addressed earlier in this chapter, is due to the two interpretations that may be given scores in the intermediate range: one, that intermediate scores are due to a confused understanding of the subtest; and two, that intermediate scores are due to a conception of the aspect of scientific theories not adequately represented by the subtest alternatives.

One way of investigating the ambiguity of the intermediate-range scores is to interview students whose subscores are in this range. Interviews could be used to probe student understanding of the subtest items. Discrimination between the two different interpretations of intermediate-range subtests could then be made using the specific information obtained from student interviews. It is conceivable, of course, that within a given population both interpretations would be required to adequately explain all intermediate range scores. However, identification of a prevalent misconception in a particular population (such as preservice elementary teachers) would suggest that a revised version of the COST should incorporate the identified misconception. Inclusion of more than two alternative conceptions in a revised version of the COST would require a different item format (e.g., multiple choice) and an appropriate scoring procedure.

A second type of research possibility elicited by the study consists of the following potential uses of the COST:

1. Determining teachers' understanding of the tentative and revisionary conception of the nature of science.
2. Discriminating and describing teachers' conceptions of particular aspects of scientific theories.
3. Assessing student outcomes in college educational programs whose goal is to teach a particular conception of the nature of science.
4. Investigation of factors associated with the use of inquiry teaching strategies and inquiry-oriented programs.

The first two of the previously listed potential uses of the COST are concerned with determining teachers' conceptions of the nature of science. This use of the COST is of considerable importance in research on the effectiveness of teacher education programs. For example, in the previous section of this chapter it was suggested that appropriately designed instruction in the development of scientific knowledge would lead to preservice teachers' understanding of the tentative and revisionary conception of science. The COST would be useful in determining the effectiveness of such an instructional program. Additionally, the COST would be useful in assessing outcomes in any educational program whose goal is understanding the particular aspects of scientific theories addressed by the COST.

Of particular importance to this investigator is research on the use of inquiry teaching strategies and inquiry-oriented programs. As discussed previously in this dissertation, the importance of a teacher's conception of the nature of science as a potentially significant influence on his/her teaching behavior requires a means of determining that conception. The COST, an instrument organized around an educationally and socially significant conception of science, was developed in response to that need. Consequently, the COST is especially relevant to research on factors associated with the use of particular teaching strategies and programs.

Even though "inquiry teaching" is fraught with ambiguity (Shulman & Tamir, 1973, pp. 1111-1116), the use of particular variations of this theme is emphasized in programs which teach "science as inquiry" (Stake & Easely, 1978, pp. 2-4). Investigations of factors that influence successful inquiry teaching should certainly assess teachers' understanding of the tentative and revisionary conception of science, a conception of science intimately related to understanding "science as inquiry."

APPENDICES

APPENDIX A

PILOT FORM A OF THE COST

APPENDIX A

PILOT FORM A OF THE COST

Name______________________

Student No.________________

Instructions: For each statement indicate the extent of your agreement by circling the appropriate number and then marking that number on the answer sheet. Use the following scale:

Strongly agree	Agree	Disagree	Strongly disagree
(1)	(2)	(3)	(4)

For example: The moon is made of cheese.

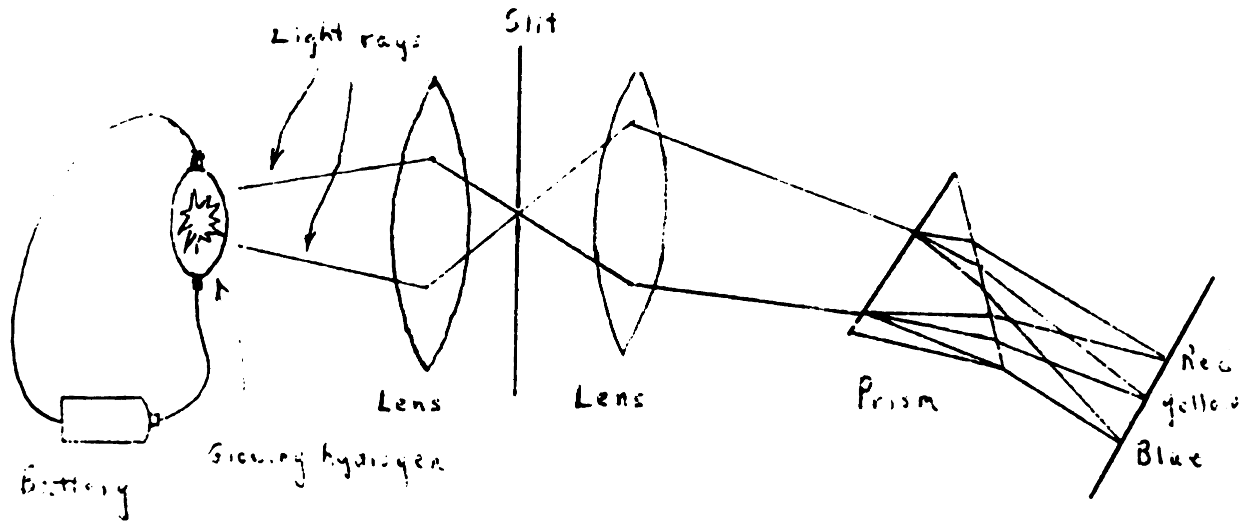
1	2	3	(4)

Then mark 4 at the appropriate place on the answer sheet.

Make sure you give a response for every statement!

TOUT I

During the last quarter of the 19th century Balmer investigated the pattern of light (spectrum) that results when light from hot, glowing hydrogen is passed through a prism (refer to diagram below). He observed a regularity in the spacing of the distinct colors that made up the spectrum.



Balmer accounted for the spacing between colors by applying a mathematical formula that he developed. Even though he could use the formula to calculate the spacing between colors, he could not explain why the spectrum occurred.

In 1913, Niels Bohr published a theory of the atom. His theory was based on the study of the spectrum of hydrogen and could be used to explain why that spectrum occurred. Bohr's theory described the atom as consisting of a nucleus surrounded by orbiting electrons that are found particular distances from the nucleus. The fact that orbits only occurred at particular distances from the nucleus could be related to the observation that the hydrogen spectrum consisted of lines of light only found at particular places.

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

1. The atom as described by Bohr must be regarded as existing because his theory is supported by the evidence.

   1 2 3 4

2. If Bohr's theory is correct, there should be enough evidence to prove it conclusively.

   1 2 3 4

3. The success of Bohr's theory in explaining the atom depends on the method he used to develop his theory.

   1 2 3 4

4. In choosing between Bohr's theory and the more recent theory that replaced it, it was possible to make the choice by comparing them against the facts.

   1 2 3 4

5. Bohr's theory is to be judged a successful scientific theory only if it can be used in explaining the spectra of other elements than hydrogen.

   1 2 3 4

6. Bohr's theory has been used to make predictions. If the predictions are found to be correct, then this proves Bohr's theory.

   1 2 3 4

7. Even if Bohr's theory is correct it will never be proved conclusively.

   1 2 3 4

8. It doesn't matter how Bohr developed his theory as long as it explains evidence that concerns the atom.

   1 2 3 4

9. Bohr's theory is only a way of organizing scientists' observations. It doesn't make any claims about what is actually there.

   1 2 3 4

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

10. Bohr's theory was not derived from his observations of the hydrogen spectrum, but invented in order to account for them.

    1 2 3 4

11. For Bohr's theory to be called a scientific theory, it must be capable of predicting things that can be observed.

    1 2 3 4

12. Even though the hydrogen atom cannot be observed, if it is assumed that Bohr's theory is correct, then it must be a true description of the hydrogen atom.

    1 2 3 4

13. Bohr's theory may be judged a successful scientific theory if all it does is show why Balmer's formula is correct.

    1 2 3 4

14. Because Bohr's theory was replaced by another theory in 1925, that more recent theory must be closer to the truth.

    1 2 3 4

15. We have no basis for claiming that the theory of the atom that replaced Bohr's theory is a better approximation to the truth.

    1 2 3 4

TOUT II

In 1859, Charles Darwin published his theory of biological evolution. This theory proposed that all living things change and that the plants and animals living today were not the first plants and animals. Darwin also proposed that the mechanism of evolution was natural selection. The process of natural selection, according to Darwin's theory, led to the survival of those individuals in a population who were best adapted to their environments. These individuals preferentially passed their characteristics on to future generations.

The clues that led Darwin to his theory of evolution were several. Some of them are summarized below:

a. knowledge of geology that included an awareness of the tremendous age of the earth and the idea that the geologic features of the earth had changed over time.

b. the diversity of closely related varieties of organisms that lived on the Galapagos Islands.

c. studies of variation due to artificial selection in the breeds of domestic pigeons.

d. knowledge that the potential for population growth ultimately exceeds the capacity of the environment required for it.

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

16. Darwin's theory of evolution would be capable of predicting and explaining phenomena that were not known when Darwin developed the theory.

    1 2 3 4

17. Darwin's theory was deliberately devised for directing scientific research and for finding connections between things in the natural world that would otherwise be regarded as unrelated.

    1 2 3 4

18. If Darwin's ideas on evolution couldn't be tested, they wouldn't be part of a scientific theory.

    1 2 3 4

19. If scientists wish to resolve the conflict Darwin's theory has with another scientific theory, they should compare them to the facts.

    1 2 3 4

20. Because Darwin's theory is a scientific theory, it will never be proved conclusively.

    1 2 3 4

21. Because Darwin's theory of evolution is supported by the evidence, we should recognize that "natural selection" is a process that exists in the natural world.

    1 2 3 4

22. In order to develop his theory, Darwin used a set of scientific rules for developing theories from data.

    1 2 3 4

23. One reason the controversy between Darwin's theory of evolution and the creationist theory of life cannot be settled is because the advocates of each theory interpret the data according to their own theory.

    1 2 3 4

Strongly Agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

24. It is legitimate for those who don't accept Darwin's theory to wait until it is conclusively proved before accepting it.

    1 2 3 4

25. Because Darwin's theory is considered to be correct, it must be a description of the natural world as it actually exists.

    1 2 3 4

26. Darwin didn't use an established scientific method to develop his theory from the observed facts. He invented his theory in order to account for the facts.

    1 2 3 4

27. Other theories have been used to explain how evolution occurs, but none has been as useful as Darwin's. Because several theories have been used to explain evolution, "natural selection" may only be considered a useful idea and may not be claimed to be a process that exists.

    1 2 3 4

28. In preparation for developing his theory, Darwin collected as much data as possible. He must have done this without any preconceived ideas in order to maintain scientific objectivity.

    1 2 3 4

29. Darwin's theory merely added to the knowledge of evolution that existed before he developed his theory.

    1 2 3 4

30. Indicate the extent of your agreement with the following argument:

    If Darwin's theory is correct, then B should be observed.
    B has been observed.
    Therefore, Darwin's theory is correct.

    1 2 3 4

TOUT III

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

31. The following is an example of a scientific theory: John's shoes are wet and muddy. He must have walked through a rain puddle.

 1 2 3 4

32. A scientific theory is a system of related statements.

 1 2 3 4

The following types of arguments might be used to test a scientific theory. Assuming that no errors have been made in the observations, indicate your agreement with the arguments.

33. If the theory is correct, we should observe X. We have observed X. Therefore, the theory is proved.

 1 2 3 4

34. If the theory is correct, we should observe X. We have observed X. Therefore, the theory has some support.

 1 2 3 4

35. How a scientific theory is generated is irrelevant to its usefulness.

 1 2 3 4

36. Science provides us with methods that when used according to the rules lead us from observed facts to theories.

 1 2 3 4

37. When a scientific theory is well supported by evidence, the objects postulated by the theory must be regarded as existing.

 1 2 3 4

38. A correct scientific theory is a true description of reality.

 1 2 3 4

39. There are no pure observations in our world. All observations are influenced by our ideas.

 1 2 3 4

40. When two theories are available to explain the same range of natural phenomena, choice may be made between the two theories using an objective, scientific procedure.

 1 2 3 4

APPENDIX B

PILOT FORM B OF THE COST

APPENDIX B

PILOT FORM B OF THE COST

Name______________________________

Student No.________________________

<u>Instructions</u>: For each statement indicate the extent of your agreement by circling the appropriate number and then marking that number on the answer sheet. Use the following scale:

Strongly agree	Agree	Disagree	Strongly disagree
(1)	(2)	(3)	(4)

For example: The moon is made of cheese.

1	2	3	(4)

Then mark 4 at the appropriate place on the answer sheet.

Make sure you give a response for every statement!

TOUT IV

Geologists have accumulated evidence that leads them to claim that about two hundred and fifty million years ago glaciers covered parts of what are now South America, Antarctica, India, Africa, and Australia. During this time there were no glaciers of any kind in the northern continents.

In the 1930's geologists developed a theory which postulated that some continents were connected by land bridges called isthmian links. The presence of the isthmian links could be used to explain the occurrence of weather patterns that gave rise to the distribution of glaciers two hundred and fifty million years ago.

In the 1960's the theory of plate tectonics, which was completely at odds with the old theory, was developed by Dietz and Hess. This theory describes the surface of the earth as consisting of huge plates that move about constantly. The theory is used to explain the glaciation of two hundred and fifty million years ago as well as a wide variety of other geological phenomena. In fact many geologists cite the glaciation of 250 million years ago as proof of the theory of plate tectonics.

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

1. Because the postulated isthmian links existed millions of years ago and, therefore, can't be observed, it is not appropriate to claim either that they existed or didn't exist.

   1 2 3 4

2. Dietz and Hess didn't invent the theory of plate tectonics. They objectively derived it from the facts.

   1 2 3 4

3. Dietz and Hess invented their theory.

   1 2 3 4

4. If plate tectonics is a legitimate scientific theory, it must have testable implications.

   1 2 3 4

5. In order to determine the value of plate tectonics we should know what method Dietz and Hess used to develop the theory.

   1 2 3 4

6. The theory of plate tectonics should be capable of explaining and predicting phenomena that were not known when the theory was developed.

   1 2 3 4

In testing the theory of plate tectonics, it was hypothesized that the thickness of sediment on the ocean bottom should increase the further one sampled from the Mid-Atlantic Ridge (a geologic formation that is the boundary between two plates). Sampling was done, and the hypothesis was confirmed. (Use this information in responding to #7 and #8.)

7. This proves the theory.

   1 2 3 4

8. This only provides support for the theory.

   1 2 3 4

9. Claims concerning the existence or nonexistence of the isthmian links can be made depending on the evidence.

   1 2 3 4

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

10. Plate tectonics is closer to the truth than the theory of isthmian links.

    1 2 3 4

11. Plate tectonics is a new theory. Given enough time it's likely that enough evidence will be accumulated to prove it conclusively.

    1 2 3 4

12. Even though no one ever saw the isthmian links, if there had been enough evidence in support of them, we could claim that they actually existed.

    1 2 3 4

13. If the theory of plate tectonics were only a guess, it would still qualify as a scientific theory.

    1 2 3 4

The glaciation of 250 million years ago has been used as evidence in support of two completely different theories. Indicate your agreement with the following two explanations of this contradiction.

14. Choice of conflicting theories must depend on something else than objective observation.

    1 2 3 4

15. Some of the scientists involved in this research must have made a mistake.

    1 2 3 4

TOUT V

Questions about the origin of life are some of the most interesting a human being can ask. In 1938, a Russian bio-chemist, A. I. Oparin, proposed a theory to explain the origin of life. He argued that the atmosphere of the earth before the origin of life was very different from what it is today. Under the conditions of this early atmosphere, Oparin claimed that simple molecules came together to form more complex organic substances that are the constituents of living systems. Eventually, according to the theory, the organic substances combined together to form more and more complex substances, until a structure formed that we would call living.

Since Oparin developed his theory many experiments have been done to test it. In 1953, Stanley Miller published a paper that described his attempts to test some of the claims of Oparin's theory. Miller simulated conditions that were thought to duplicate those of the earth's early atmosphere. Under these conditions he was able to produce many complex organic substances that are found in living organisms.

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

16. Several theories have been proposed to explain the origin of life. Because of this, no matter what the evidence, we can never consider anyone of them to be a description of what actually happened.

    1 2 3 4

17. Oparin's ideas wouldn't be considered a scientific theory unless they were useful in guiding research into the origin of life.

    1 2 3 4

18. The success of Oparin's theory depends on the methods he used to develop it.

    1 2 3 4

19. The ultimate test of Oparin's theory is whether enough evidence can be found to prove it conclusively.

    1 2 3 4

20. Oparin's theory will be discarded if it's disproved by the facts.

    1 2 3 4

21. The events explained by Oparin's theory can't be observed because they happened billions of years ago. Therefore, no matter how many experiments are done that support this theory, it can never be known that the events his theory described actually happened.

    1 2 3 4

22. It has been proposed that life on earth was created by a supernatural event. This is not a scientific theory because by its very nature it is not subject to experimental investigation.

    1 2 3 4

23. Oparin's theory of the origin of life will not be discarded until another theory replaces it.

    1 2 3 4

24. Any facts that are used in explanations of how life may have originated are interpreted in terms of some theory.

    1 2 3 4

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

25. It is possible that someday enough evidence will have been accumulated in support of Oparin's theory so that scientists will be justified in saying: Yes! This theory is true! The events it describes actually happened.

    1 2 3 4

26. Science is open-minded! Therefore, any guess concerning the origin of life can be considered a scientific theory.

    1 2 3 4

27. In judging the value of Oparin's theory to the scientific community it is important for us to know how he developed his theory.

    1 2 3 4

28. Oparin must have used an established, scientific method to develop his theory.

    1 2 3 4

An argument similar to the following could be used to describe Miller's experiments:

A. If organic substances that serve as the basis of life were formed in the earth's early atmosphere, then a simulation of the conditions thought to exist in the early atmosphere should give rise to organic substances that might serve as the basis of life.

B. Organic substances that might serve as the basis of life appeared under the simulated conditions.

C. Therefore, organic substances that might serve as the basis of life formed in the earth's early atmosphere.

29. This argument proves Oparin's theory.

    1 2 3 4

30. This argument doesn't prove Oparin's theory. It does provide some support for it.

    1 2 3 4

TOUT VI

Strongly agree	Agree	Disagree	Strongly disagree
(1)	(2)	(3)	(4)

31. Observation is not a basis for evaluating scientific theories because of the influence of theory on observation.

    1 2 3 4

32. A successful scientific theory offers a unified account of different phenomena.

    1 2 3 4

33. Scientific theories must have observable implications.

    1 2 3 4

34. Good scientific theories must eventually be proved conclusively.

    1 2 3 4

35. When two scientific theories are available to explain the same range of natural phenomena, it is not possible to compare them against the observed facts.

    1 2 3 4

36. Because scientific theories are primarily conceptual tools for organizing experience, the unobservable objects postulated by some theories cannot be claimed to exist.

    1 2 3 4

37. A scientific theory may never be proved conclusively.

    1 2 3 4

38. The usefulness of a scientific theory depends on the methods used to derive the theory from the facts.

    1 2 3 4

39. By applying a scientific method to their data, scientists develop theories.

    1 2 3 4

40. The acceptance of scientific theories does not commit us to the existence of the things postulated by the theories.

    1 2 3 4

APPENDIX C

FINAL FORM OF THE COST

APPENDIX C

FINAL FORM OF THE COST

Starting Time____________________

Finishing Time____________________

This questionnaire is intended to assess your conceptions of various aspects of scientific theories. There are no right answers to any of the items. However, in order to get an accurate description of your conceptions, it is important that you think carefully about every item.

The items are organized around several scientific theories. Each set of items is prefaced by a brief description that provides information about the scientific theory relevant to the set. You may use that information in answering items. You may find that not all the information you need to answer an item is available. In that case, do the best you can drawing on your understanding of the item. It is presumed you have little or no understanding of the theories addressed by the item.

Instructions: For each item indicate the extent of your agreement by choosing the appropriate category and then marking the number appropriate to that category on the answer sheet. Use the following scale:

Strongly agree	Agree	Disagree	Strongly disagree
(1)	(2)	(3)	(4)

Make sure you give a response for every statement!

Geological Theories

Geologists have accumulated evidence that leads them to claim that about two hundred and fifty million years ago glaciers covered parts of what are now South America, Antarctica, India, Africa, and Australia. During this time there were no glaciers of any kind in the northern hemisphere.

In the 1930's geologists developed a theory which postulated that some continents were connected by land bridges called isthmian links. The presence of the isthmian links could be used to explain the occurrence of weather patterns that gave rise to the distribution of glaciers two hundred and fifty million years ago.

In the 1960's the theory of plate tectonics, which was completely at odds with the old theory, was developed independently by Dietz and Hess. This theory describes the surface of the earth as consisting of huge plates that move about constantly. The theory is used to explain the glaciation of two hundred and fifty million years ago as well as a wide variety of other geological phenomena. In fact, many geologists cite the glaciation of 250 million years ago as evidence in support of the theory of plate tectonics.

Strongly agree	Agree	Disagree	Strongly disagree
(1)	(2)	(3)	(4)

1. Because the postulated isthmian links existed millions of years ago and, therefore, can't be observed, it is not appropriate to claim either that they existed or didn't exist.

   1 2 3 4

2. Dietz and Hess didn't invent the theory of plate tectonics. They objectively derived it from the facts.

   1 2 3 4

3. In order to determine the value of plate tectonics we should know what method Dietz and Hess used to develop the theory.

   1 2 3 4

In testing the theory of plate tectonics, it was hypothesized that the thickness of sediment on the ocean bottom should increase the further one sampled from the Mid-Atlantic Ridge (a geologic formation that is the boundary between two plates). Sampling was done, and the hypothesis was confirmed.

4. This proves the theory.

   1 2 3 4

5. Plate tectonics is a new theory. Given enough time it's likely that enough evidence will be accumulated to prove it conclusively.

   1 2 3 4

6. Even though no one ever saw the isthmian links, if there were enough evidence in support of them, we could claim that they actually existed.

   1 2 3 4

Evidence of the glaciation of 250 million years ago has been used as support for two completely different theories. Indicate your agreement with the following two explanations of this contradiction.

7. Choice of conflicting theories must depend on something else than objective observation.

   1 2 3 4

8. Some of the scientists involved in this research must have made a mistake.

   1 2 3 4

Oparin's Theory of Abiogenesis

In 1938, a Russian bio-chemist, A. I. Oparin, proposed a theory to explain the origin of life. He argued that the atmosphere of the earth before the origin of life was very different from what it is today. Under the conditions of this early atmosphere, Oparin claimed that simple molecules came together to form more complex organic substances that are the constituents of living systems. Eventually, according to the theory, the organic substances combined together to form more and more complex substances, until a living structure was formed.

Since Oparin developed his theory many experiments have been done to test it. In 1953, Stanley Miller published a paper that described his attempts to test some of the claims of Oparin's theory. Miller simulated conditions that were thought to duplicate those of the earth's early atmosphere. Under these conditions he was able to produce many complex substances that are constituents of living organisms.

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

9. Several theories have been proposed to explain the origin of life. Because of this, no matter what the evidence, we can never consider anyone of them to be a description of what actually happened.

   1 2 3 4

10. The success of Oparin's theory depends on the methods he used to develop it.

    1 2 3 4

11. The ultimate test of Oparin's theory is whether enough evidence can be found to prove it conclusively.

    1 2 3 4

12. Oparin's theory of the origin of life will not be discarded until another theory replaces it.

    1 2 3 4

13. Any facts that are used in explanations of how life may have originated are interpreted in terms of some theory.

    1 2 3 4

14. If Oparin's theory is ever accepted scientists will then be justified in saying: Yes! This theory is true! The events it describes actually happened.

    1 2 3 4

15. In judging the value of Oparin's theory to the scientific community it is important for us to know how he developed his theory.

    1 2 3 4

An argument similar to the following could be used to describe Miller's experiments:

A. If organic substances that serve as the basis of life were formed in the earth's early atmosphere, then a simulation of the conditions thought to exist in the early atmosphere should give rise to organic substances that might serve as the basis of life.

B. Organic substances that might serve as the basis of life appeared under the simulated conditions.

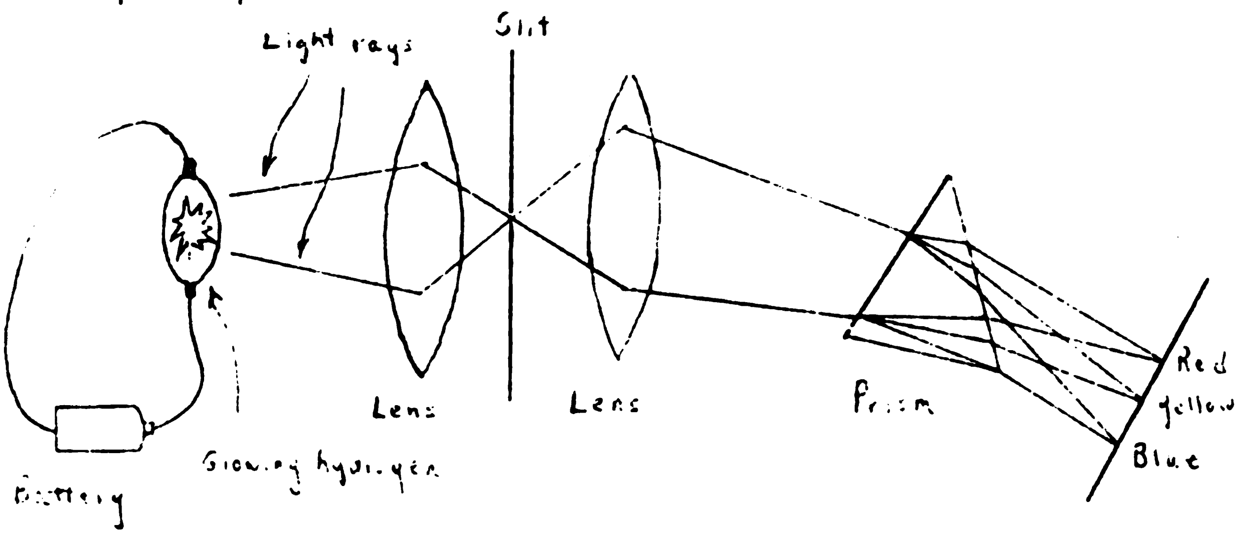
C. Therefore, organic substances that might serve as the basis of life were formed in the earth's early atmosphere.

16. This argument proves Oparin's theory.

    1 2 3 4

Bohr's Theory of the Atom

During the last quarter of the 19th century Balmer investigated the pattern of light (spectrum) that results when light from hot, glowing hydrogen is passed through a prism (refer to diagram below). He observed a regularity in the spacing of the distinct colors that made up the spectrum.



Balmer accounted for the spacing between colors by applying a mathematical formula that he developed. Even though he could use the formula to calculate the spacing between colors, he could not explain why the spectrum occurred.

In 1913, Niels Bohr published a theory of the atom. His theory was based on the study of the spectrum of hydrogen and could be used to explain why that spectrum occurred. Bohr's theory described the atom as consisting of a nucleus surrounded by orbiting electrons that are found particular distances from the nucleus. The fact that orbits only occurred at particular distances from the nucleus could be related to the observation that the hydrogen spectrum consisted of lines of light only found at particular places.

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

17. The atom as described by Bohr must be regarded as existing because his theory is supported by the evidence.

    1 2 3 4

18. If Bohr's theory is correct, there should be enough evidence to prove it conclusively.

    1 2 3 4

19. Even if Bohr's theory is correct it may never be proved conclusively.

    1 2 3 4

20. It doesn't matter how Bohr developed his theory as long as it explains evidence that concerns the atom.

    1 2 3 4

21. Bohr's theory was not derived from his observations of the hydrogen spectrum, but invented in order to account for them.

    1 2 3 4

22. Even though the hydrogen atom cannot be observed, if it is assumed that Bohr's theory is correct, then it must be a true description of the hydrogen atom.

    1 2 3 4

23. Because Bohr's theory was replaced by another theory in 1925, that more recent theory must be closer to the truth.

    1 2 3 4

24. We have no basis for claiming that the theory of the atom that replaced Bohr's theory is a better approximation to the truth.

    1 2 3 4

Darwin's Theory of Evolution

In 1859, Charles Darwin published his theory of biological evolution. This theory proposed that all living things change and that the plants and animals living today were not the first plants and animals. Darwin also proposed that the mechanism of evolution was natural selection. The process of natural selection, according to Darwin's theory, led to the survival of those individuals in a population who were best adapted to their environments. These individuals preferentially passed their characteristics on to future generations.

The clues that led Darwin to his theory of evolution were several. Some of them are summarized below:

a. knowledge of geology that included an awareness of the tremendous age of the earth and the idea that the geologic features of the earth had changed over time.

b. the diversity of closely related varieties of organisms that lived on the Galapagos Islands.

c. studies of variation due to artificial selection in the breeds of domestic pigeons.

d. knowledge that the potential for population growth ultimately exceeds the capacity of the environment to provide for it.

Strongly agree (1)	Agree (2)	Disagree (3)	Strongly disagree (4)

25. If scientists wish to resolve the conflict Darwin's theory has with another scientific theory, they should compare them to the facts.

1 2 3 4

26. Because Darwin's theory of evolution is supported by the evidence, we should recognize that "natural selection" is a process that exists in the natural world.

1 2 3 4

27. One reason the controversy between Darwin's theory of evolution and the creationist theory of life cannot be settled is because the advocates of each theory interpret the data according to their own theory.

1 2 3 4

28. It is legitimate for those who don't accept Darwin's theory to wait until it is conclusively proved before accepting it.

1 2 3 4

29. Darwin didn't use an established scientific method to develop his theory from the observed facts. He invented his theory in order to account for the facts.

1 2 3 4

30. Other theories have been used to explain how evolution occurs, but none has been as useful as Darwin's. Because several theories have been used to explain evolution, "natural selection" may only be considered a useful idea and may not be claimed to be a process that exists.

1 2 3 4

31. In preparation for developing his theory, Darwin collected as much data as possible. He must have done this without any preconceived ideas in order to maintain scientific objectivity.

1 2 3 4

32. Indicate the extent of your agreement with the following argument:

If Darwin's theory is correct, then B should be observed.
B has been observed.
Therefore, Darwin's theory is correct.

1 2 3 4

General Questions on Scientific Theories

Strongly agree	Agree	Disagree	Strongly disagree
(1)	(2)	(3)	(4)

The following type of agrument might be used to test a scientific theory. Assuming that no errors have been made in the observations, indicate your agreement with the argument.

33. If the theory is correct, we should observe X. We have observed X. Therefore, the theory is proved correct.

1 2 3 4

34. How a scientific theory is generated is irrelevant to its usefulness.

1 2 3 4

35. When a scientific theory is well supported by evidence, the objects postulated by the theory must be regarded as existing.

1 2 3 4

36. When two theories are available to explain the same range of natural phenomena, choice may be made between the two theories using an objective, scientific procedure.

1 2 3 4

37. Observation is not a basis for evaluating scientific theories because of the influence of theory on observation.

1 2 3 4

38. Scientific theories may eventually be proved conclusively.

1 2 3 4

39. By applying a scientific method to their data, scientists develop theories.

1 2 3 4

40. The acceptance of scientific theories does not commit us to the existence of the things postulated by the theories.

1 2 3 4

Respondent Information

Answers to the following questions will be used in interpreting your responses to the items on the questionnaire.

Use the following scale to describe the state of your knowledge of the following subjects before you took this questionnaire.

	State of Knowledge				
Subject	Mastery	Highly Competent	Somewhat Competent	Slightly Competent	No Knowledge
41. Bohr's theory of the atom	1	2	3	4	5
42. Theory of plate tectonics	1	2	3	4	5
43. Theories of abiogenesis	1	2	3	4	5
44. Darwin's theory of evolution	1	2	3	4	5
45. Philosophy of science	1	2	3	4	5

Estimate to what extent your responses to the following sets of items were influenced by personal convictions independent of your understanding of the subjects.

	Influence of Personal Conviction				
Item Set	Complete	Strong	Moderate	Weak	None
46. Bohr's theory of the atom	1	2	3	4	5
47. Geological theories	1	2	3	4	5
48. Darwin's theory of evolution	1	2	3	4	5
49. Theory of abiogenesis	1	2	3	4	5

50. How would you rate the effort you made in answering the items on this questionnaire?

considerable	moderate	very little	none
(1)	(2)	(3)	(4)

APPENDIX D

MULTI-TRAIT AND MULTI-METHOD MATRIX

Pearson Correlation Coefficients

	BGEN	BTES	BTHC	DONT	DGEN	DTES	DTHC	GONT	GGEN	GTES	GTHC	AONT	AGEN	ATES	ATHC	NONT	NGEN	NTES	NTHC
BONT	.1059 (110) S=.135	.1487 (110) S=.060	.1797 (110) S=.030	.4451 (110) S=.001	.0519 (110) S=.295	.2687 (110) S=.002	.0852 (110) S=.188	.1129 (110) S=.120	.2053 (110) S=.016	.3132 (110) S=.001	-.0014 (110) S=.494	.1494 (110) S=.060	-.0222 (110) S=.409	.2367 (110) S=.006	-.1424 (110) S=.069	.4828 (110) S=.001	.0197 (110) S=.419	.3573 (110) S=.001	-.0449 (110) S=.321
BGEN		.1854 (110) S=.025	-.1329 (110) S=.083	.1300 (110) S=.088	.2518 (110) S=.004	.2185 (110) S=.011	.1136 (110) S=.119	.1077 (110) S=.131	.3194 (110) S=.001	.1695 (110) S=.038	.0571 (110) S=.277	.0313 (110) S=.373	.2971 (110) S=.001	.2041 (110) S=.016	.0237 (110) S=.403	.0612 (110) S=.263	.2835 (110) S=.001	.2676 (110) S=.002	.0459 (110) S=.317
BTES			.1324 (110) S=.084	.0842 (110) S=.191	.0824 (110) S=.196	.2756 (110) S=.002	.1983 (110) S=.019	-.0725 (110) S=.226	.1287 (110) S=.090	.4361 (110) S=.001	.1106 (110) S=.125	.1519 (110) S=.057	.1285 (110) S=.384	.4253 (110) S=.001	-.0378 (110) S=.348	.3355 (110) S=.001	.0095 (110) S=.461	.2553 (110) S=.004	.0970 (110) S=.157
BTHC				.0110 (110) S=.455	-.2042 (110) S=.016	.1861 (110) S=.026	.1099 (110) S=.127	.0525 (110) S=.293	-.0598 (110) S=.267	.1367 (110) S=.077	.0653 (110) S=.249	.0284 (110) S=.384	-.1655 (110) S=.042	.0353 (110) S=.357	.0039 (110) S=.484	.1794 (110) S=.030	-.0429 (110) S=.328	.0239 (110) S=.402	.2180 (110) S=.011
DONT					.1345 (110) S=.081	.4011 (110) S=.001	.2111 (110) S=.013	.1774 (110) S=.032	.1185 (110) S=.109	.1563 (110) S=.052	.0214 (110) S=.412	.0935 (110) S=.166	-.0819 (110) S=.198	.2562 (110) S=.003	.0143 (110) S=.441	.3000 (110) S=.001	.0379 (110) S=.347	.3670 (110) S=.001	.2098 (110) S=.014
DGEN						-.0730 (110) S=.224	-.1458 (110) S=.064	.1796 (110) S=.030	.0992 (110) S=.151	-.0804 (110) S=.202	-.1073 (110) S=.132	-.0126 (110) S=.448	.2440 (110) S=.005	-.1127 (110) S=.120	-.0260 (110) S=.394	.0652 (110) S=.249	.0791 (110) S=.206	.1589 (110) S=.049	.1014 (110) S=.146
DTES							.1132 (110) S=.119	.0621 (110) S=.260	.2397 (110) S=.006	.2975 (110) S=.001	.0394 (110) S=.341	.0396 (110) S=.341	.0967 (110) S=.158	.3733 (110) S=.001	.0082 (110) S=.466	.2557 (110) S=.004	.2846 (110) S=.001	.4911 (110) S=.001	.2158 (110) S=.012
DTHC								.0941 (110) S=.164	-.1252 (110) S=.096	.1111 (110) S=.124	.2176 (110) S=.011	.2624 (110) S=.003	-.1996 (110) S=.018	.2558 (110) S=.009	.2978 (110) S=.001	.0859 (110) S=.186	.0779 (110) S=.209	.0621 (110) S=.260	.2683 (110) S=.002
GONT									-.1066 (110) S=.134	.0215 (110) S=.412	.1307 (110) S=.087	.1942 (110) S=.021	-.0246 (110) S=.399	.1372 (110) S=.076	.0699 (110) S=.234	-.0268 (110) S=.390	.0585 (110) S=.272	.2024 (110) S=.017	.2121 (110) S=.013
GGEN										.2890 (110) S=.001	-.2024 (110) S=.017	-.0650 (110) S=.250	.4188 (110) S=.001	.2957 (110) S=.001	-.0767 (110) S=.213	.0975 (110) S=.155	.3175 (110) S=.001	.1868 (110) S=.025	.0642 (110) S=.253
GTES											.2084 (110) S=.014	.1713 (110) S=.037	.1149 (110) S=.116	.5310 (110) S=.001	.0434 (110) S=.326	.2922 (110) S=.001	.0645 (110) S=.251	.2628 (110) S=.003	-.0645 (110) S=.252
GTHC												.0624 (110) S=.259	-.1871 (110) S=.025	.0801 (110) S=.203	.1303 (110) S=.087	-.0084 (110) S=.465	-.0252 (110) S=.397	.0645 (110) S=.252	-.0353 (110) S=.357
AONT													-.1620 (110) S=.045	.3051 (110) S=.001	.2250 (110) S=.009	.0293 (110) S=.381	.1827 (110) S=.028	.1098 (110) S=.127	.1227 (110) S=.101
AGEN														.1608 (110) S=.047	-.0979 (110) S=.154	.0880 (110) S=.180	.2982 (110) S=.001	.0978 (110) S=.155	.1753 (110) S=.033
ATES															.0482 (110) S=.309	.3148 (110) S=.001	.1739 (110) S=.035	.2996 (110) S=.001	.1567 (110) S=.051
ATHC																-.2175 (110) S=.011	.4447 (110) S=.001	.0115 (110) S=.452	.1540 (110) S=.054
NONT																	-.0628 (110) S=.257	.3169 (110) S=.001	.1561 (110) S=.052
NGEN																		.1458 (110) S=.064	.3233 (110) S=.001
NTES																			.2250 (110) S=.009

REFERENCES

REFERENCES

Aikenhead, G. The measurement of knowledge about science and scientists: An investigation into the development of instruments for formative evaluation. Unpublished doctoral dissertation, Harvard University, 1972.

Aikenhead, G. The measurement of high school students' knowledge about science and scientists. Science Education, 1973, 57(4).

Bates, G. A search for subscales in the Science Process Inventory. Paper presented at the annual meeting of the National Association for Research in Science Teaching, 1974.

Borich, G. D., Malitz, D., & Kugle, C. L. Convergent and discriminant validity of five classroom observation systems: Testing a model. Journal of Educational Psychology, 1978, 70(2), 119-128.

Bridgham, R. G. Concepts of science and learning science. School Sciences, November 1969.

Brody, B. A. (Ed.). Readings in the philosophy of science. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1970.

Campbell, D. T. Reforms as experiments. American Psychologist, April 1969, 218-223.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56(2), 81-105.

Cooley, W. W., & Klopfer, L. Test on Understanding Science (Form W). Princeton, N.J.: Educational Testing Service, 1961.

Cronbach, L. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16(3), 297-332.

Cronbach, L., & Meehl, P. E. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 281-302.

Dampier, Sir W. C. A history of science (4th ed.). Cambridge, England: Cambridge University Press, 1958.

Einstein, A., & Infeld, L. The evolution of physics. New York: Simon & Schuster, 1938.

Hanson, N. R. Patterns of discovery. London: Cambridge University Press, 1965.

Hempel, C. G. Philosophy of natural science. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Hillis, S. R. The development of an instrument to determine student views of the tentativeness of science. In Research and curriculum development in science education: Science teacher behavior and student affective and cognitive learning (Vol. 3). Austin, Texas: University of Texas Press, 1975.

Hungerford, H., & Walding, H. The modification of elementary methods students' concepts concerning science and scientists. Paper presented to the 1974 convention of the National Science Teachers Association, 1974.

Hurd, P. D. Science literacy: Its meaning for American schools. Educational Leadership, 1958, 16, 13-16.

Johnson, P. G. The goals of science education. Theory Into Practice, 1962, 1(5), 239-244.

Jungwirth, E. Testing for understanding of the nature of science. Journal of College Science Teaching, February 1974, 206-210.

Kimball, M. E. Understanding the nature of science. Journal of Research in Science Teaching, 1967-68, 5(2), 110-120.

Korth, W. Test every senior project: Understanding the social aspects of science. Paper presented at the 42nd meeting of the National Association for Research on Science Teaching, 1969.

Kuhn, T. S. The structure of scientific revolutions (2nd ed.). Chicago: The University of Chicago Press, 1970a.

Kuhn, T. S. Reflections on my critics. In I. Lakatos & A. Musgrave (Eds.), Criticism and the growth of knowledge. London: Cambridge University Press, 1970b.

Kuhn, T. S. The essential tension. Chicago: University of Chicago Press, 1977.

Laudan, L. Progress and its problems: Toward a theory of scientific growth. Berkeley: University of California Press, 1977.

Long, A., & Murray, C. The 1980 federal R & D budget. Chemical and Engineering News, January 29, 1979, 19-24.

Lucas, A. M. Hidden assumptions in measures of knowledge about science and scientists. Science Education, 1975, 59(4), 481-485.

Magnusson, D. Test theory. Reading, Mass.: Addison-Wesley Publishing Co., 1967.

Martin, M. Concepts of science education. Glenview, Ill.: Scott, Foresman & Co., 1972.

Morabito, L. A., Synnott, S. P., Kupferman, P. N., & Collins, S. A. Discovery of currently active extraterrestrial volcanism. Science, 1979, 204(4936).

Nagel, E. The structure of science. New York: Harcourt, Brace, & World, Inc., 1961.

National Science Teachers Association. The NSTA position on curriculum development in science. The Science Teacher, 1962, 29(9), 32-37.

National Science Teachers Association. NSTA position statement on school science education for the 70's. The Science Teacher, 1971, 38(8), 46-47.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, O. Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill Book Co., 1975.

Pella, M. O., O'Hearn, G. T., & Gale, C. W. Referents to scientific literacy. Journal of Research in Science Teaching, 1966, 4.

Ravetz, J. R. Scientific knowledge and its social problems. New York: Oxford University Press, 1973.

Robinson, J. T. The nature of science and science teaching. Belmont, Calif.: Wadsworth, 1968.

Rubba, P. Nature of Scientific Knowledge Scale. Bloomington, Ind.: Indiana University, 1976.

Schwab, J. J. The teaching of science as inquiry. In J. J. Schwab & P. F. Brandwein (Eds.), The teaching of science. Cambridge, Mass.: Harvard University Press, 1962.

Scientific Literacy Research Center. Wisconsin Inventory of Science Processes. Madison, Wis.: The University of Wisconsin, 1967.

Shamos, M. The price of scientific literacy. _National Association of Secondary School Principals_, 1963, _47_, 41-51.

Shulman, L., & Tamir, P. Research on teaching in the natural sciences. In R. N. W. Travers (Ed.), _Second handbook of research on teaching_. Chicago: Rand McNally College Publishing Co., 1973.

Stake, R. E., & Easely, J. _Case studies in science education_ (Vol. 1). Washington, D.C.: U.S. Government Printing Office, Stock No. 038-000-00377-1.

Stegmüller, W. _The structure and dynamics of theories_. New York: Springer-Verlag, 1976.

Stice, G. _Facts About Science Test_. Princeton, N.J.: Educational Testing Service, 1958.

Weiss, I. _Report of the 1977 national survey of science, mathematics, and social studies education_. Washington, D.C.: U.S. Government Printing Office, Stock No. 038-000-00364-0.

Welch, W. W. _Welch Science Process Inventory_ (Form D). Minneapolis, Minn.: Author, 1969.

Yager, R. E. Priorities for research in science education: A study committee report. _Journal of Research in Science Teaching_, 1978, _15_(2).