AN EXPERIMENTAL ANALYSIS OF STATIC VISUAL ACUITY OF NOVICE, TRAINED AND EXPERIENCED MOTORCYCLISTS DURING SIMULATED MOTORCYCLE OPERATION

DISSERTATION
Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of the Ohio State University by Gary Lee Winn

* * * * * * * * * * * *
Ohio State University
1985

Reading Committee:
Dr. Marlin Languis
Dr. William Loadman
Dr. William Umstattd

Approved by:
Dr. William Umstattd
Adviser
Department of Educational Theory and Practice
Dedication

For her endless inspiration and encouragement, the author dedicates the final document to Sherry A. Winn.
Acknowledgements

The author would like to call attention to the contributions made by the following people:

Dr. William Umstattd, Ohio State University, for tireless assistance and counseling in the preparation of this document;

Dr. Edwin Novak, Ohio State University, for guidance in the early stages of the dissertation research;

Dr. James Aaron, Southern Illinois University, for providing initiative and encouragement;

Mr. Robin Tomhagen Harrison, for providing encouragement and boundless enthusiasm;

Mr. Robert Vorak, for help in creating a computer program that made data collection a reliable task;

Mr. Larry Bignall, for providing a class of trained novice motorcyclists to serve as subjects;

Mr. S. K. Lee, for valuable assistance in reducing experimental data and help in data analysis;

Mr. Robert Arroyo, who provided editing and clerical assistance in preparation of the final document;

Mrs. Trisha Miller, who provided valuable clerical assistance in preparing early document drafts;

The staff of the American Motorcyclist Association, many of whom volunteered as experimental subjects.
Vita

JANUARY 21, 1949 . . . . . born, San Antonio, Texas

EDUCATION

June 10, 1967 . . . . . High School, Walter E. Stebbins, Dayton, Ohio
June 7, 1971 . . . . . Bachelor of Arts, Wright State University, Dayton, Ohio
April 15, 1975 . . . . . Master of Arts, University of Dayton, Dayton, Ohio

WORK EXPERIENCE


July, 1973-April, 1975 . . . Graduate Assistant, University of Dayton, Psychology Dept., Dayton, Ohio

September, 1975-June, 1976 . Graduate Assistant, University of Houston, Psychology Dept., Houston, Texas

September, 1976-August, 1985. Manager, Legislative Affairs, American Motorcyclist Association, Westerville, Ohio
MEMBERSHIPS AND HONORARIES

1978-1985 . . . . Society of Automotive Engineers
1979-1985 . . . . National Committee on Uniform Traffic Laws and Ordinances
1980-1985 . . . . Transportation Research Board
1975 . . . . . . Psy Chi, an honorary in psychology
1984 . . . . . . Epsilon Pi Tau, an honorary in technology
1984 . . . . . . Phi Kappa Phi, a scholastic honorary

PUBLICATIONS


# TABLE OF CONTENTS

Dedication .................................................. ii
Acknowledgements ............................................ iii
Vita ............................................................ iv
Memberships and Honoraries ................................. v
Publications ..................................................... v
Table of Contents ........................................... vi
List of Tables .................................................. ix
List of Figures ................................................ xi
Chapter I. INTRODUCTION .................................... 1
  Background of Study ........................................ 1
  Problem Statement .......................................... 6
  Hypotheses .................................................. 7
  Research Assumptions ...................................... 9
  Research Delimitations ..................................... 11
  Research Limitations ...................................... 12
  Definitions .................................................. 13
  Significance of the Study .................................. 18
  Use of the Study's Results ................................ 19

Chapter II. LITERATURE REVIEW .......................... 20
  The Issue of Motorcycle Accidents ...................... 20
  Motorcycle Accident Characteristics ................... 26
  Motorcycle Accident Analysis ............................ 26
  Motorcycle Accident Countermeasures ................... 36
  Automobile and Motorcycle Operation ................... 43
  Perception Processes and Motor Vehicle Operation ...... 45
# Table of Contents, Continued

## Chapter II, continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Processing: Adopting a Paradigm</td>
<td>48</td>
</tr>
<tr>
<td>Developing Visual Skills in Drivers</td>
<td>62</td>
</tr>
<tr>
<td>Motorcycle Training in the United States and Canada</td>
<td>63</td>
</tr>
<tr>
<td>Summary and Critique of Relevant Literature</td>
<td>65</td>
</tr>
</tbody>
</table>

## Chapter III. RESEARCH PROCEDURES AND METHODS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Use of Simulation</td>
<td>70</td>
</tr>
<tr>
<td>Subjects</td>
<td>72</td>
</tr>
<tr>
<td>Visual Task Apparatus</td>
<td>74</td>
</tr>
<tr>
<td>Motor Task Apparatus</td>
<td>75</td>
</tr>
<tr>
<td>Stimuli</td>
<td>76</td>
</tr>
<tr>
<td>Research Design</td>
<td>79</td>
</tr>
<tr>
<td>Variables</td>
<td>81</td>
</tr>
<tr>
<td>Control Groups</td>
<td>82</td>
</tr>
<tr>
<td>Test Procedures and Instructions</td>
<td>82</td>
</tr>
<tr>
<td>The Pilot Study</td>
<td>84</td>
</tr>
<tr>
<td>Pilot Study Results</td>
<td>86</td>
</tr>
<tr>
<td>Summary of Research Procedures and Methods</td>
<td>88</td>
</tr>
</tbody>
</table>

## Chapter IV. RESEARCH RESULTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics of Test Subjects</td>
<td>91</td>
</tr>
<tr>
<td>Validity Analysis</td>
<td>93</td>
</tr>
<tr>
<td>Item Analysis</td>
<td>94</td>
</tr>
<tr>
<td>Test Reliability</td>
<td>105</td>
</tr>
<tr>
<td>Task Learning</td>
<td>106</td>
</tr>
<tr>
<td>Group Means</td>
<td>111</td>
</tr>
<tr>
<td>Assumptions of the Analysis of Variance</td>
<td>113</td>
</tr>
<tr>
<td>Gender Differences</td>
<td>120</td>
</tr>
<tr>
<td>Multivariate Results</td>
<td>123</td>
</tr>
<tr>
<td>The Analysis of Variance</td>
<td>124</td>
</tr>
<tr>
<td>Visual Task ANOVA Results</td>
<td>125</td>
</tr>
<tr>
<td>Motor Task ANOVA Results</td>
<td>127</td>
</tr>
<tr>
<td>Reaction Time Task ANOVA Results</td>
<td>129</td>
</tr>
<tr>
<td>Between Groups Comparison</td>
<td>131</td>
</tr>
<tr>
<td>Correct vs. Incorrect Reaction Times</td>
<td>132</td>
</tr>
<tr>
<td>Discussion of Research Results</td>
<td>133</td>
</tr>
</tbody>
</table>
# Table of Contents, Continued

Chapter V. CONCLUSIONS AND RECOMMENDATIONS ........................................... 140
  Review of the Problem ...................................................................................... 140
  Review of the Rationale of the Study .............................................................. 141
  Review of the Methods and Procedures .......................................................... 142
  Review of Research Findings ......................................................................... 143
  Conclusions from the Study's Results ............................................................. 145
  General Recommendations from the Study's Results ..................................... 148
  Developing a Paradigm ...................................................................................... 152
  Interpreting the Research Conclusions and Recommendations ..................... 153

APPENDIX ............................................................................................................ 155
  A. Computer Program for Apparatus Timers ................................................. 155
  B. Schematic of Electronic Interface ............................................................... 158
  C. Schematic of Experimental Apparatus ....................................................... 159
  D. Letter, Number, Arrow, and Landolt Ring Stimuli ..................................... 160
  E. Scoring Sheet ............................................................................................... 161
  F. Subject Consent Form (Ohio State Form HS-027) ....................................... 162

LIST OF REFERENCES .......................................................................................... 163
List of Tables

2. Experience in Accident-Involved Motorcycle ................. 31
3. Accident-Involved Motorcyclist Training ..................... 32
4. Typical Human Errors in Equipment Operation and Their Possible Causal Factors .......... 59
5. Demographics of Test Subjects .................................. 92
7. Item Analysis: Proportions of Visual Task Items Correct by Subject Group .................. 97
8. Item Analysis: Proportions of Motor Task Items Correct by Subject Group ..................... 99
9. Comparison of Mean Correct Responses, First Ten Versus Last Ten Items, All Groups .......... 107
10. Comparing First Ten Scores to Last Ten by Group ................ 109
11. Mean Task Scores by Subject Group .......................... 112
12. Variance on Reaction Time Task, by Group .................... 114
13. Subject Error Scores by Subject Group ......................... 119
14. Comparing Male to Female Task Scores, by Group ............... 121
15. Comparing Male to Female Task Scores, by Group (Probability Levels) ....................... 122
16. Multivariate Summary Table .................................... 124
17. Univariate Summary Table, Visual Task Errors ............... 126
18. Univariate Summary Table, Motor Task Errors .................. 128
List of Tables, continued.

19. Univariate Summary Table, Reaction Time . . . . . . . . 130
20. Pairwise Comparisons for Reaction Time Task . . . . . . .
21. Differences between Correct and Incorrect Responses, by Group . . . . . . . . . . . . . . . . . . . . . . 133
List Of Figures

1. States with Self-Funding Rider Education Legislation ............. 4
3. Collision Contact Points ........................................ 30
4. Interface of the Vehicle, Driver and the Environment ............... 37
5. Relationship of Stimulus Similarity, Response Similarity, and Transfer of Training .......... 55
7. Pilot Study Results ............................................ 88
Chapter I

INTRODUCTION

There is scarcely a problem in the United States that commands more media attention than the problem of highway safety. The National Highway Traffic Safety Administration reports that the traffic death toll exceeds 45,000 lives each year, eclipsed only by certain diseases as the major public health problem. As a subset, motorcycle accidents and deaths take a chilling toll on Americans, largely male youth entering their most productive years. Motorcyclists in the 16-24 year-old category have a fatality rate between 3 and 4 times worse than their four-wheeled counterparts, according to the National Safety Council (1983).

Solutions to the problems of motorcycle safety have ranged from banning motorcycles, to helmet use laws, to highly restrictive licensing requirements. In general, these solutions have been unpopular among affected groups and more often assailed for their limited statistical effect.

Background of Study

In the last decade, motorcycle rider groups have been joined by major motorcycle manufacturers, and more recently,
the United States Department of Transportation, in advocating motorcycle operator skill training to reduce accident and fatality rates. Many people view education as the most cost-effective and politically feasible method of lowering morbidity rates.

The appeal for motorcycle operator training, especially to educators, remains intuitive due to a lack of empirical evidence that shows motorcycle skills training actually leads to reduction in accidents or improvement in skills. Despite the widespread appeal for the education of motorcycle users, there needs to be conclusive evidence on whether motorcycle training works or not. This would be followed by the development of an explanatory paradigm to guide future research in motorcycle safety problems.

The tide is rising rapidly in favor of increasing motorcycle training funds and facilities with or without evidence that training is effective. Legislators in eighteen states have enacted legislation in cooperation with motorcycle users to semi-permanently fund rider training activities. Ohio may pass such legislation in 1985. Indeed, the stakes are not small as indicated by the number of states that have embraced training programs, and the degree of public expenditures on training that may or may not work. For example, Illinois
annually takes in 1.2 million dollars from increases in motorcycle registration fees, and places the money in a fund dedicated to support motorcycle training activities. Ohio would take in even more money. Over 1.4 million dollars would be appropriated if House Bill 291 becomes law in the 116th session of the General Assembly. Figure 1 shows the current number of states with this type of "self-funding" education legislation.
Figure 1.
States With Self-Funding Rider Education Legislation.
If motorcycle skill training programs serve their intended purpose, they might be expanded and continued indefinitely. Conversely, if they are ineffective, they should be abolished as a grand and honorable scheme but unnecessary for further legislative enactments or embrace by the traffic safety community. Yet, shouldn't all these decisions be guided by research?

It is possible that motorcycle education could be accomplished through existing automobile training programs, which would save project development costs. The literature strongly underscores that the untrained novice motorcyclist makes inappropriate responses to hazards for which auto training would not prepare the motorcycle operator.

How, then, do novice motorcycle riders develop into experienced, accident-free operators? The answer remains elusive. Perhaps motorcycle operator training consists of the opportunity to experience a host of novel situations under controlled conditions. For example, the novice motorcyclist must deal with new controls, such as a hand throttle and brake, manual or hand clutch, and novel seating position. In addition, there are roll, pitch, and yaw characteristics.
It takes years for an experienced motorcyclist to learn to adapt successfully to these controls and environmental conditions, and still maneuver safely. Formal instruction for novices could shorten the time it takes to accommodate new controls in traffic. Training could become a way to in-still experience artificially, and do it much more rapidly than the 3 or 4 years it takes to become an "experienced rider" (Hurt, 1981).

Problem Statement

In 1984, there were more than 5.6 million motorcycles registered in the United States. During that year, there were 172,000 reported motorcycle accidents, and 4,500 motorcyclist fatalities. The accident rates were worse for those riders in the 16-24 age group than in any other group. Research also demonstrates that riders who have very little experience and no formal instruction are much more likely to be involved in an accident (Hurt, 1981).

Many of these riders make no accident avoidance maneuvers, or make inappropriate, confused responses in an emergency. Motorcycle education is one possible solution to the accident problem. In the last decade, for example, millions
of dollars have been spent by the motorcycle industry to develop standardized curricula. More recently, eighteen states passed legislation to create motorcycle education programs. Despite this, there is little empirical evidence that motorcycle training either reduces accident frequency, improves motor skills or perceptual capabilities of students.

This study compares basic perceptual and motor skills of trained riders to untrained novices, experienced riders and non-motorcyclists. The study should help fill the mentioned needs and guide curriculum development, too. This research will help create an empirical rationale for one aspect of motorcycle training, the aspect dealing with visual acuity motor skills and reaction time. Findings could have a direct impact on saving lives because valid research is necessary to support further legislation and enhance training opportunities for beginners.

Hypotheses

Hypotheses are objective criteria that allow test results to be interpreted on mathematical principles. Two sets of hypotheses are presented: the first set is designed to test the control procedure to limit the pre-test of experience with motorcycle controls and assure the experimenter
that all subjects in all groups learned the task. A task-learning analysis compares the first and last ten test items for each task - Visual, Motor, and Reaction Time. For task learning, the null hypotheses are:

a) The first task learning hypothesis states that there is no statistically significant difference between the first ten and last ten Visual Task error scores across all groups.

b) The second task learning hypothesis states that there is no statistically significant difference between the first ten and last ten Motor Task error scores across all groups.

c) The third task learning hypothesis further states that there is no statistically significant difference between the first ten and the last ten Reaction Time scores across all groups.

The second set of null hypotheses tests the implicit expectation that experienced riders will perform better than trained riders, who in turn will perform the tasks better than untrained motorcyclists, and finally, non-motorcyclists. These null hypotheses are stated thus:

d) The first null hypothesis investigated states that there is no statistically significant difference in Visual Task errors on a simulated motorcycle performance test among groups of trained, untrained, experienced riders, and non-motorcyclists.
e) The second null hypothesis investigated states that there is no statistically significant difference in Motor Task errors on a simulated motorcycle performance test among groups of trained, untrained, experienced and non-motorcyclists.

f) The third null hypothesis investigated states that there is no statistically significant difference in Reaction Time on a simulated motorcycle performance test among groups of trained, untrained, experienced, and non-motorcyclists.

g) The fourth null hypothesis investigated states that there is no statistically significant difference in Visual Task, Motor Task, and Reaction Time considered together on a simulated motorcycle performance test among groups of trained, untrained, experienced, and non-motorcyclists.

Research Assumptions

Empirical research must make certain assumptions. It presumes validity for certain research procedures or tests. Research assumptions deserve mention, because if they later prove baseless, or incorrect, the research results are open to question. These assumptions for this study are that:
1. Experimental motorcycle simulation adequately represents "real-world" motorcycle operation.

2. The visual and motor tasks simulate actual motorcycle operations adequately.

3. Experimental apparatus is calibrated properly to provide for accurate measurements of reaction time and decision-making errors.

4. Measurements do not vary appreciably during or between tests.

5. All research subjects represent themselves accurately by giving honest answers during the pre-experimental interview.

6. Reaction time and error rate are valid measures of critical motorcycle operating skills, and can discriminate among novice, trained, experienced and non-motorcyclists.

7. Measurement will be transferred to record sheets accurately.

8. Laboratory simulation can duplicate threatening traffic conditions adequately.

9. Fewer errors and shorter reaction times in simulated motorcycle operation translate into more appropriate and quicker reactions under real traffic conditions.
Research Delimitations

Research delimitations refer to those random variables over which the experimenter has a degree of control. If not controlled, these variables limit the study's generality. D'Amato (1970) lists three types of delimitations which he calls environmental, or situation variables, subject variables, and sequence variables. For this study, research delimitations are that:

1. Environmental variables (lighting, temperature, and noise) are controlled sufficiently so that subjects are tested under consistent conditions.

2. Sequence variables are controlled by randomizing the order of subjects and the order of stimuli.

3. Subject-relevant variables are controlled by stratifying groups for age, sex, and pre-test vision. All subjects except those in a non-motorcyclist category should be qualified motor vehicle operators at the time of the test.

4. A standard set of instructions should be used.

5. A standard length for practice should be given to each subject.

6. Subjects should be tested at as near as possible the same time of day to control hunger and fatigue effects.
Research Limitations

Research involving humans is subject to limitations, or those factors over which the experimenter lacks direct control. These may influence results. Some research limitations that may influence the study's results are that:

1. It is not possible to classify all subjects precisely as "experienced", "trained", "novice" riders or "non-motorcyclists".

2. Levels of individual differences among subjects (intelligence quotients, level of alertness and special skills) are variables which could account for observed differences.

3. A motorcycle simulator is limited in the things it can do that simulate real-life; for example, it cannot duplicate traffic noise, temperature, weather conditions, motorcycle balance and threatening traffic conditions. Thus, a simulator may be easier to operate than an actual motorcycle.

4. The experimenter will not have control over subjects who plan to participate, but do not because they are ill or move away. That is to say, samples may be reduced by circumstances beyond the control of the experimenter.
5. The results of a laboratory study may not generalize to a field situation.

Definitions

All terms used in research must have an operating definition that can be specified through a testing operation that provides a criterion for its application (Hempel, 1966). The operant definitions related to this study follow:

* Acuity: The ability to resolve detail visually in an environmental field.

* Alpha-numeric stimulus: A scale of one-character, least-confused alphabetic and numeric characters used as visual targets; M, B, T, L, 1, 3, 5, and 7. There are 8 alpha-numeric visual targets. (See Kinney, Marsetta, and Showman, 1966) Alphanumeric stimuli are presented in Helvetica medium, 30 point type.

* Arrow: A visual stimulus whose orientation is up, down, left, or right. When viewing an arrow stimulus, subjects indicate its orientation verbally. Arrow stimuli are presented in Helvetica medium, 30 point type.
* Errors: An incorrect motor or verbal response made by a subject.

* Experienced Motorcycle Rider (ER): Refers to self-taught or informally taught motorcycle operators with forty-eight or more months of riding experience, or more than 10,000 miles. The operator must be free of reported accidents in the last thirty-six months. Full current motorcycle endorsement is necessary. This rider has never had a formal training course.

* Formally Trained Novice (FTN): Refers to a recently formally taught motorcyclist within the first full season of operation. "Taught" refers to a course of standard instruction presented by certified instructional personnel; the course must consist of fifteen or more hours of education. These riders are eligible as experimental subjects between one and twelve months after completion of a formal motorcycle education course, and must be free of reported accidents in that time.
* **Landolt Ring:** The Landolt ring is a circle, or ring, open in one quadrant. It was designed in 1906 as an alternative to the problem of Latin alphabet literacy and currently as an alternative to the Snellen chart. It was also used to test glare recovery time during night vision examinations. The gap can be open at the top, bottom, right, or left, and as a stimulus, the task is to verbally identify the gap orientation. All Landolt ring stimuli are presented in Helvetica medium, 30 point type.

* **Motorcycle:** In Ohio, motorcycle means "every motor vehicle other than a tractor having a saddle for the use of the operator, and designed to travel on not more than three wheels in contact with the ground" (Traffic Laws Annotated, National Committee on Traffic Laws and Ordinances, 1972). This definition does not include motor vehicles commonly known as motor-driven cycles, mopeds, or motor scooters.

* **Motorcycle Controls:** Refers generally to the specific set of hand-controlled devices such as levers,
switches, signals and prop stands that are peculiar to starting, stopping and operating a motorcycle.

* Motor Task: This experimental task simulates a portion of actual motorcycle operation, specifically the front hand brake and left-handed horn button, clutch lever and throttle.

* Non-Motorcyclist (NM): The non-motorcyclist sample comes from the population at large, including drivers and non-drivers of automobiles; non-motorcyclists have no experience in operating a motorcycle.

* Novice Motorcyclist (NOV): This person has limited experience operating a motorcycle, ranging from a few minutes to a season. He or she is not formally trained, is legally licensed to operate a motor vehicle in Ohio, and must be free of reported accidents.

* Reaction Time (RT): Defined as the difference in hundredths of a second between the stimulus slides onset and the time the response is made on the motorcycle simulator.

* Simulator: This apparatus consists of the handlebar of an actual motorcycle with levers, switches,
cables, throttle and handgrips. It is mounted to a wooden platform and has adjustments for height and reach.

* Time-sharing: "Time-sharing refers to situations in which a human being has two or more chores to which he has to alternate his attention. In a strict sense, an individual cannot give simultaneous attention to two or more aspects of a situation" (McCormick, 1970).

* Visual Field: As reported by Burg in 1966, the average visual field is about 175 degrees for 16-55 year-olds (males and females averaged).

* Visual Task: The visual task is a defined, randomly presented series of alphabetic or numeric stimuli or Landolt rings or arrows (see Appendix D). Each of these targets in the visual task are commonly used in acuity research (see Burg, 1964; Neil, Sampson, and Gribben, 1971; Markowitz and Weitzman, 1968), and the subject is required to verbally name the letter or number, or give the orientation of the arrow or Landolt ring.
Significance of the Study

The public will continue to demand accountability in highway safety programs and in spending public funds. Yet, despite the outcry for improvements in motorcycle safety and expenditures by private industry in the development of standardized curricula for motorcycle operator education, answers remain elusive for the obvious questions, "Does motorcyclist education work?" and "What can be done to make the instruction better?".

If education for motorcyclists works, it can be expanded and supported further in schools and in the legislature because it saves lives. If it does not work, it is unworthy of continued public embrace, expenditures of monies, or legislative involvement. Applied research can answer these questions and guide further research. In addition, any reliable apparatus for the experiment can be made available for future researchers and for replication work. This would include a full motorcycle handlebar simulator with actual motorcycle controls; a minicomputer to time reactions, record responses for appropriateness; an electronic interface to link the minicomputer and simulator and electronic switching and stimulus exposure equipment.
Use of the Study's Results

The study's results should be used to provide substance and data related to the public's concerns in highway safety. These data:

1. Will confirm or refute that trained motorcyclists perform a motorcycle simulation task better than untrained riders.

2. Should suggest whether motorcyclists can be tested for visual acuity adequately on a simulator.

3. Could support further curriculum development activities focusing on psychomotor skill development.


5. Can support future legislation providing funds, manpower and organization to motorcyclist education activities.

6. Can be integrated into a larger framework of human factors knowledge, and provide clues about all human performance.

7. Could support the development of operator-licensing station equipment to screen out underqualified motorcyclists.
Chapter II

LITERATURE REVIEW

A review of the literature related to proposed research can guide its development and design, identify and correct shortcomings of past studies, and save time for the researcher by eliminating redundant investigations. In this chapter, literature is examined that illustrates the problems in motorcycle safety and how historical and current attempts have been made to deal with them.

Basic human perceptive processes and a comparison of motorcycle and automotive operation are noted in this chapter. Overall, a framework is constructed so that the present study is supported for relevancy, timeliness, and for use of established and accepted procedures.

The Issue of Motorcycle Accidents

An extensive review of available historical literature on the aspects of motorcycle safety unfortunately revealed little of value from which to draw conclusions. A review of popular motorcycle publications from 1926 to 1948 did not mention motorcycle accident problem in the overall highway safety problem of those decades.
The National Safety Council has statistics available on automobile safety beginning in 1913; its statistics for motorcycles began in 1963 (NSC, 1983), leaving a large void. The National Highway Traffic Safety Administration (NHTSA), the highway safety branch of the Federal Department of Transportation, has not established long-term trends in motorcycle safety; it is unclear how far back NHTSA's motorcycle safety statistics go historically, but probably to about 1968, when the agency was organized. The American Motorcyclist Association (AMA) has verified state motorcycle accident data back to 1963 (Winn, 1977), and because it coincides with the starting point of data of the NSC, it is logical to begin examining trends in that year by using AMA's statistics.

In 1963, the AMA records 769,237 motorcycle registrations, which have risen to 5,500,000 in 1983, an increase of over 600 percent. During this same period, auto registrations increased from 83 million to 163 million, an increase of only 94 percent (NSC, 1983). The upward swing in motorcycle registrations in the mid-1960s suggests their increased popularity, perhaps owing to improved reliability and selection brought about primarily by Japanese imports in the past two decades. Sharp increases in motorcycle registrations
can again be seen in the late 1970's, coinciding with sharp world increases in fuel prices.

Along with increased registrations, motorcycle accidents have risen from 24,000 to 172,000 between 1963 and 1983, an increase of 617 percent (AMA, 1984). Motorcyclist fatalities have risen from 728 across all states in 1963 to 4,500, an increase of 518 percent; a trend almost identical to the accident increase. Figure 2 overlays motorcycle registration, accident and fatality trends over these two decades.
Figure 2.
Source: American Motorcyclist Association.
These descriptive statistics are gross indicators. There is good evidence that annual reports of motorcycle accidents may be underestimates. Krause, Franti, and Riggins (1973) show that "57% of the 7,699 cases identified as serious injury cases had not been reported to the police agencies" during an in-depth study of motorcycle injuries in California.

The National Highway Traffic Safety Administration suggested the trend in motorcycle accidents may continue to rise during the next decade (U.S. Department of Transportation, 1983). Assuming flat market trends, the government agency performed a linear regression on existing data, "the result being a projection of about 6,700 motorcyclist deaths in 1990" (U.S. DOT, 1983). This represents a jump of almost 50% in a single decade and, if accurate, is certainly a cause for alarm among highway safety officials. Table 1 illustrates NHTSA's projections of fatality rates for all classes of vehicle occupant until 1990.
Table 1

Vehicle Occupant Fatalities

<table>
<thead>
<tr>
<th>VEHICLE</th>
<th>Vehicles in Use (1)</th>
<th>Fatality Rate (2)</th>
<th>Fatalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger Cars</td>
<td>106.9(3)</td>
<td>120.0(3)</td>
<td>21.7</td>
</tr>
<tr>
<td>Light Trucks &amp; Vans</td>
<td>31.0(3)</td>
<td>41.5(4)</td>
<td>20.6</td>
</tr>
<tr>
<td>Medium &amp; Heavy Trucks</td>
<td>5.7</td>
<td>6.4(4)</td>
<td>17.4</td>
</tr>
<tr>
<td>Motorcycles</td>
<td>6.0</td>
<td>7.5</td>
<td>71.7</td>
</tr>
<tr>
<td>TOTALS</td>
<td>149.6(5)</td>
<td>175.4(5)</td>
<td>23.3</td>
</tr>
</tbody>
</table>

(1) Millions
(2) Occupant fatalities per 100,000 vehicles
(3) Based on R. L. Polk data
(5) Excludes buses

Source: Department of Transportation; Traffic Safety Trends and Forecasts, 1983.

NHSTA (1983) shows that the occupant fatality rates in passenger cars to be about 21.7 deaths per 100,000 vehicles registered. By comparison, the motorcycle rate is listed by the agency as 71.7, about 3 1/2 times higher.
Motorcycle Accident Characteristics

Unfortunately, most motorcycle accidents involve injury to the operators and their passengers. Hurt (1981) shows that 45.1% of the accident victims he studied had injuries described as moderate or worse; only 2.5% of the motorcycle user accident cases had no injury at all. The National Accident Sampling System (NASS, 1983) reports that 24% of the injured operators have injuries requiring hospitalization, compared to only 3% of automobile and truck operators. Even if no further details were provided, this tragic backdrop of motorcycle accidents, injuries and fatalities in perspective to all types of motor vehicle accidents would demand efforts to remedy the problem. A few studies are available to further identify the problems in motorcycle safety.

Motorcycle Accident Analysis

No nationwide study has been accomplished that examines motorcycle accidents in depth. The logistics and costs would be overwhelming, especially when sampling practices show that a reliable and valid sample has many of the characteristics of the parent population. One such study of high regard has been carried out in southern California under contract to
the NHTSA; the 1977-1980 investigation of 900 on-scene accidents in the Los Angeles area (Hurt, Oullette, and Thom, 1981) contributed much to the understanding of the dynamics of United States motorcycle crashes despite its regional character. Hurt's teams reconstructed motorcycle accidents and interviewed participants to depict critical man-machine events. Two significant findings of the study follow:

1. Most of these accidents investigated occurred at intersections, where scanning skills and quick reflexes are critical; 66.7% of all these mishaps occurred at crossing or "T" intersections, or alleys and driveways. This finding by Hurt closely replicates Harano and Peck (1968), (who found a similarly high incidence of intersection accidents for motorcycle operators).

2. Even though the "other vehicle" operator was at fault in over 50% of the crashes, the motorcycle operator contributed little to accident prevention. Apparently, many or the accidents actually caused by another vehicle driver could have been prevented if the motorcyclist made appropriate responses or evasive maneuvers. No evasive action of any sort was made by the involved motorcyclist in 31.9% of Hurt's investigated accidents, as astonishing as that seems. Inappropriate action occurred in 76.2% of the accidents examined by Hurt. For example, most motor-
cycles have two independent brakes; it is important to know how to use both brakes. However, riders in Hurt's study only used the rear brake in 30.2% of the cases and neither brake in another 31.4% of the cases, thus ignoring the far greater stopping ability of the front brake. Riders correctly used both brakes (the front hand brake and the rear foot brake) only on 17% of all occasions.

Hurt, et al., noted, "It is important that the motorcycle rider must detect, decide and react to a traffic hazard in less than two seconds (on average). The proper evasive action must be taken and executed well without any delay." Yet, for some reason, in 3 of 4 cases, the proper actions were not taken. Since most operators in Hurt's study were untrained, and since actions to avoid an accident were inappropriate in 3 of 4 cases, it seems that even rudimentary training could benefit the average motorcyclist.

Hurt found that few riders used horns to attempt to prevent any accident. Ninety-three percent did not use the horn because they either did not know where the control was, felt it was unnecessary, or had no time to react.

Most accidents occurred to the front of the motorcycle, that is, directly in the line of sight. Figure 3 shows that
in 61.9% of the cases studied by Hurt, the pending collision occurred directly in front, or to the front left or the front right of the machine. NHTSA's nationwide Fatal Accident Reporting System (U. S. Department of Transportation, 1983) showed that 75.1% of motorcycle multi-vehicle accidents and 62.2% of single motorcycle accidents occurred to the rider's front visual area, thus corroborating Hurt on a national study of fatal-only motorcycle accidents. Polanis (1979) reported similar findings that accidents usually happen to the rider's front visual area. See Figure 3.
Figure 3.
Collision Contact Points.
Source: Hurt, Ouellet and Thom, 1981.
The part played by training and experience is of tremendous significance in the accidents studied by Hurt. Of the riders in their sixth or less month of motorcycle use, 57.4% were involved in accidents on that particular motorcycle. Of those in the 10-twelve month category, the figure rose to 73.3%, compared to only 3.2% of the accidents happening to those with over four years of experience. Table 2 presents Hurt's findings on motorcyclist experience.

Table 2

Experience on Accident-Involved Motorcycle

<table>
<thead>
<tr>
<th>Category Label</th>
<th>Absolute Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-6 Months</td>
<td>491</td>
<td>54.6</td>
<td>57.5</td>
</tr>
<tr>
<td>7-12 Months</td>
<td>136</td>
<td>15.1</td>
<td>73.3</td>
</tr>
<tr>
<td>1-2 Years</td>
<td>112</td>
<td>12.4</td>
<td>86.4</td>
</tr>
<tr>
<td>2-3 Years</td>
<td>63</td>
<td>7.0</td>
<td>93.8</td>
</tr>
<tr>
<td>3-4 Years</td>
<td>26</td>
<td>2.9</td>
<td>96.8</td>
</tr>
<tr>
<td>Over 4 Years</td>
<td>27</td>
<td>3.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Unknown</td>
<td>45</td>
<td>5.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: Hurt, 1981, p. 129

Ninety-two percent of the operators in Hurt's sample had no formal training; by far, most were self-taught.
Table 3 presents Hurt's findings on training.

Table 3

<table>
<thead>
<tr>
<th>Category Label</th>
<th>Absolute Frequency</th>
<th>Relative Frequency</th>
<th>Cumulative Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Taught</td>
<td>400</td>
<td>44.4</td>
<td>49.5</td>
</tr>
<tr>
<td>Family - Friends</td>
<td>343</td>
<td>39.1</td>
<td>92.0</td>
</tr>
<tr>
<td>Motorcycle course</td>
<td>41</td>
<td>4.6</td>
<td>97.0</td>
</tr>
<tr>
<td>By Professionals</td>
<td>20</td>
<td>2.2</td>
<td>99.5</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>.4</td>
<td>100.0</td>
</tr>
<tr>
<td>Unknown</td>
<td>92</td>
<td>10.2</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: Hurt, 1981, p. 115

Hurt suggests a human factors solution to one problem in motorcycle braking. Interconnected motorcycle brakes might circumvent the important problem of over-braking the rear brake and under-braking the front brake, and as long as the rear brake pedal is used, both brakes would be applied. Hurt notes that one particular brand of factory-equipped motorcycle with integrated brakes was unfortunately too underrepresented in the sample to make inferences. The author’s implication about interconnected brakes is clear, however. Operators, and especially novices learning about motorcycle controls, clearly are at risk if they
cannot remember which lever to use, or which way the throttle turns, for example.

Integrated brakes would simplify the choice problem, especially for novices. The Hurt data suggest a novice simply may be "overloaded" with task requirements and fail to execute any move properly. He summarizes by saying, "Training in collision avoidance, braking, cornering and traffic accident strategy are sure to reduce traffic involvement" (p. 153).

Some theories of motorcycle accident causes are specifically rejected by the southern California study. Some proponents argue the idea that larger-engine machines are involved in more motorcycle accidents. That is, the larger and more powerful the engine, the more likely the accident. The NHTSA, among others, has been an historical advocate of this position, but after examining the data, Hurt found no significant relationship between horsepower and motorcycle accidents. Comparing engine displacement to the appropriate exposure statistics (the relative frequency in the population at risk), Hurt discovered that 23.7% of the exposure-related motorcyclists operated engine families in the 750 cubic centimeter (cc) range, but only 15.7% of the accidents happened in that range. Conversely, the 250cc capacity motorcycles accounted for 8.3% of the exposure motorcycles but
13.3% of the accidents.

The Traffic Research Section of New Zealand's Ministry of Transport has found similar results, showing that the "engine capacity of the motorcycle (has) no significant effect on the motorcycle accident rate." (Hull, 1981) These two studies seem to suggest that in loss-of-control accidents, one major theoretical cause, engine displacement, can safely be ruled out. Hurt and Hull findings suggest an eventual human factors solution to the motorcycle safety problem, and not a hardware or simple legislative solution—banning large motorcycles.

Kaestner (1981) compares accident involved motorcyclists to auto drivers and finds major differences. Motorcycle drivers involved in accidents were younger and they more often make "head-on, other caused" accidents. Motorcyclists were 4 1/2 times more likely to sustain a fatal or serious injury and were less likely to be licensed properly for motorcycle operation. Kaestner did not examine training or experience levels in his study of Oregon drivers and motorcyclists.

Blauw and Poll (1980) showed that when contrasting the skills of experienced and inexperienced operators, the novices make certain types of errors on their motorcycles, such as lane changing, which required a long time in skill
development. The authors say, "Riding a motorcycle is a difficult task and one for which the necessary skills can be obtained only gradually." This literature search didn't reveal any other studies that compared motorcycle novices and experienced riders on variables critical to motorcycle accident prevention.

The point of reviewing the accident data is to find trouble spots and begin to locate areas of problem correction. Accident-involved motorcycle operators often have accidents that could be avoided, and many of them cause injury. Accidents happen to their front, and mostly without training, and many to inexperienced operators who worked improper or no evasive action. Chief among causes of motorcycle accidents are:

* The "other driver"
* Inexperienced
* Youthful operators
* Intersection locations
* Lack of evasive action
* Improper evasive action

While Hurt and others suggest that skill training can lead to reduced accident rates for motorcyclists, questions remain whether training leads to improved visual acuity, scanning or shorter reaction times. A further review of the
literature is called for, including current motorcycle safety countermeasures and a review of human factors processes involved in motor vehicle operation.

**Motorcycle Accident Countermeasures**

To begin a discussion of countermeasures in motorcycle safety, we should examine Figure 4. This illustration, borrowed from Crabb (1980), shows the relationship among the three basic elements of the driver situation: the driver, the environment, and the vehicle. The "hardware advocates" seem to concentrate on the vehicle-environment interface. The human factors school, on the other hand, focuses on both the driver-vehicle and the driver-environment interfaces. While each of the three possible interfaces is involved in safe motor vehicle operation, those who choose accident prevention will attempt to influence the latter two pairs. Those who choose injury mitigation or energy attenuation after a crash has occurred will choose to influence the vehicle-environment pair.
In motorcycle safety, three decades of research have been dedicated primarily to the study of the hardware available to the motorcyclist, normally helmets (Hurt, 1981; Henderson, 1975), but also improved lighting (Olson, et al., 1979; Freedman, 1982) and vehicle dynamics (Wier, 1973). Until about 1975, little was being investigated in the
motorcycle driver-environment interaction, or the motorcycle driver-vehicle interaction. After all this time, one or the other group ought to be able to show convincingly that education or hardware works to reduce accidents or attenuate energy sufficiently to improve safety without accident prevention. Neither job has been completely done.

According to Mortimer, some studies on motorcycle education have met with methodological limitations: Satten, 1980; Collins, 1979; Osga, 1982 Mortimer, 1982. Mortimer suggests that there is no particular benefit to motorcycle training based on a questionnaire study of accident rates, violations and damage costs. Both investigations were conducted on samples of riders that had completed the 20-hour course developed by the Motorcycle Safety Foundation. Osga (1982) reports that riders who pass a formal training course are more likely to have accidents; sampled course graduates had more accidents after the course. This is at odds with theories about the benefits of training.

Hurt (1981) showed that "most of the accident-involved motorcycle riders were self-taught or learned from friends or family. This group was 92% of all the accident-involved riders. After comparisons using exposure data (comparisons to the whole at-risk population), Hurt noticed that the formally trained accident rider appeared less often in accident data
than his untrained counterpart. The author says, "The trained motorcycle riders are significantly underrepresented in the accident data. The trained motorcycle rider is underrepresented in the accident data by an approximate factor of TWO" (Hurt's emphasis).

Hurt is not completely sure how the relationship between experience and accident develops, but he says in part,

"Inexperience is excessively associated with accident involvement; and inexperience is best measured by the subject (accident-involved) motorcycle. High levels of experience are underrepresented in accidents, but how is that considerable experience obtained without exposure to accidents? In the data shown, experience levels between seven months and four years clearly distinguishes that experience is beneficial. Only when the experience is much greater than four years is there a significant benefit demonstrated. It appears that specialized motorcycle rider training (author's emphasis) is the alternative which reduces risk; the acquisition of traffic experience only is simply accident exposure by comparison."

Support for motorcycle rider education based on the Hurt study alone remains weak, since it is not known for sure just what comprised "training", whether the courses were consistent or standardized, or whether they were strictly a southern California phenomenon.

Indirect evidence for motorcycle training is available through a separate California project funded by NHTSA (Anderson, Ford, and Peck, 1980). The experimental project used a three-group design comparing randomly selected riders
intending to be licensed by the state for motorcycle operation. Riders assigned to Group A were processed as though no special project was underway; they composed the control group and received the standard California licensing procedures. Group B, however, was one of the two experimental groups, and was composed of riders who received an improved licensing test, plus a three-hour "remedial on-cycle training course" in the event that the rider failed the first attempt skill test. Finally, the third group, Group C, received only the improved skill test, with no remedial training.

After a year's follow-up, the authors noted that while Groups A and C were not significantly different on measure of accidents or fatalities, Group B, with the three hour remedial course, was far lower. It may be concluded that the real benefit in accident reduction, measured at 20.3% by Anderson, et. al., was in the short training course. Because of the proper use of a control group, the results are more reliable than earlier attempts at quantifying effects of training. In addition, the authors noted full public acceptance of the improved licensing and training programs.

The authors did not do any sort of analysis that would point out which skills were critical to accident prevention.
The Anderson, Ford, and Peck study didn't address how experienced riders compared to other groups, and they didn't review the effects of the more commonly used, 20-hour training course, developed by the Motorcycle Safety Foundation.

A large-scale, federally funded field study of the efficiency of rider education is now underway in New York, again under contract to the NHTSA. The project is designed to test the feasibility of a rider education program as a condition for issuance of a license (Dwyer and McCord, 1984). The New York project is evaluating the impact of training on motorcycle operator accident rates, and the cost-effectiveness of such programs. It is due to be completed by late 1985. As the authors show, and as NHTSA explains: "Despite the promise rider education offers for reducing crashes and fatalities among motorcyclists, there have been no well-controlled evaluations of its effectiveness." An early 1985 status report revealed no clear conclusions (Dwyer and McCord, 1984).

Kenel made interesting observations about how some curriculum experts have shifted away from driver attitude research and toward a decision-making approach to accident prevention. The author illustrates the changing way researchers are looking at the problem:

"Of comparable value to curriculum development has been the results of multi-disciplinary accident investigation. True, the latter reconfirmed the age-old belief
that the driver was the major contributor in at least 85 percent of all collisions. However, assumptions con-
cerning the nature of that 85 percent changed dramat-
ically. Basically the change was a movement away from
the belief that collisions were the result of poor
attitudes and speed to an awareness that improper de-
cisions, failure to perceive critical elements and im-
proper evasive action were the major contributing causes
(emphasis by this author).
When the findings of these efforts were evaluated in
light of factors influencing perception and reports
concerning visual search patterns of drivers, it became
apparent that learning experiences were needed to im-
prove the quality of driver visual search and inform-
ation acquisition if risk assessment and responses were
to be improved." (Kenel, 1980)

The literature reviewed here shows that experts seem to
be moving away from the "hardware" school of safety and to-
wards human factors solutions which, includes motorcycle oper-
ator training. Yet, even the studies of the value of motor-
cycle training seem to have shortcomings. For example, the
studies reviewed here on the effects of rider education have
been either by questionnaire or field studies. While each
has definite scientific value, there are limitations to each.
Hays and Winkler (1971) list two problems with survey ques-
tionnaires, namely the problem of non-responses, and ques-
tion design (p. 702). With field studies there are problems
in subject-maturation, loss of subjects in the groups, and
lack of precise experimental controls (see Campbell and Stan-
ley, 1966; McGuigan, 1974). A controlled laboratory study
might offset these drawbacks, and give an idea of how
trained riders differ from non-trained riders when all other conditions are the same. A laboratory study defeats problems in non-response and question design, and since it takes place in a relatively short period of time, has no subject-maturation problems endemic in field studies. The logic for a lab study is further justified because there is no human factors literature available on whether training affects motorcycle operators differently on acuity, motor abilities, and reaction time.

Automobile and Motorcycle Operation

Any discussion of the accident kinematics of motorcycle operation is well served by a discussion of basic motorcycle engineering; what makes a motorcycle go, handle, and stop? What are the differences between motorcycle and car operation? While most citizens today operate a car, only 4 in 100 operate a motorcycle. A good definition of "motorcycle" will serve as a starting point. The National Committee on Uniform Traffic Laws (NCUTLO), defines motorcycle as follows:

"Every motor vehicle having a seat or saddle for the use of the rider and designed to travel on not more than three wheels in contact with the ground, but excluding a

Fundamentally, the differences between motorcycle and automobile operation are as follows (Motorcycle Safety Foundation, 1976):

**Brakes**..... Independent front and rear brakes, hand operated in the front and foot operated in the rear.

**Starting**..... Often foot operated, more often electric. Separate engine-stop switch for emergency use.

**Fuel**..... Fuel engaged by manual lever to turn on and off. Hand lever for choke and fuel enricher.

**Parking**..... Retractable central and side stands.

**Suspension**..... Usually adjustable front or rear for specific load.

**Signals**..... Button operated turn signals and horn.

**Power transmission**..... Foot operated lever, hand clutch.

**Handling characteristics**..... Pitch, yaw, and roll elements, separate or simultaneous.
Other...helmeted driver, gloved hands, direct contact with weather (rain, cold).

The novice motorcyclist may be overloaded when operating a motorcycle for the first time, even if the operator knows all about a car. Nevertheless, there are no requirements that a potential motorcycle buyer know how to operate a motorcycle before purchase. According to the American Motorcyclist Association (1984), three states still have no motorcycle licensing procedure which would provide a baseline requirement that novices understood the basic features of motorcycle operation.

Perception Processes and Motor Vehicle Operation

Perhaps Rockwell (1972) says it best: "If we were to design a [mechanical visual] system with the specification of the diverse skills of the human operator, most systems engineers would quickly concede defeat." Indeed, safe driving is a very complex mixture of skills, attitudes and abilities which mesh successfully in the driver with no accidents.

The process of perceiving and recognizing a signal (information) from the operator's environment is called detection, and Rockwell (1972) estimates that most of it is visual.
Detection is the process of bringing visual information from the eyes' peripheral and central visual fields, where perception occurs. The fovea of the eye is in the mid-central retina, and this is where identification occurs.

Visual acuity, or the ability to resolve detail, is best in the fovea, but motion detection is lowest in the fovea (Kaufman, 1974; Mccolgin, 1960). Looking at an object straight ahead causes signals to fall on the fovea, and foveal acuity is much better than peripheral-field acuity, whether the signal is moving or not (Klein, 1942; Green, 1970). Because of its small size, the probability of a random signal falling on the periphery of the eye is much greater than a signal falling on the 1-2 degrees of the fovea.

All possible targets cannot be identified while traveling at 80 feet per second (55 mph); some authors have postulated a selective filter to choose which peripheral signals to attend to (Mackworth and Morandi, 1967). Perhaps filter thresholds can be manipulated through training, although this has not been reported in the literature.

Target stimuli vary in their ability to be detected. They vary according to levels of brightness, contrast, luminance, and physical size. A car moving with its lights on in the peripheral field is more likely to be noticed and identified than a much smaller bicycle with no lights,
because the light provides brightness and contrast. Head movements assist in scanning the peripheral field; this is the basis of the often repeated maneuver, the "left-to-right search". Active scanning is a feature of the experienced driver (Morant and Rockwell, 1970).

Rockwell (1972) says "The role of vision in driving is believed to constitute over 90% of information input to the driver. Regardless of the exact percentage, without a doubt visual perception is paramount in vehicular control." (p. 202) Rockwell notes that "novice auto drivers, especially in their early hours of driving sample the environment close to the car," and exhibit "frantic cue searching, large eye movement travel distances, and fixations on non-relevant cues." With experience, he says, they shift to sampling strategies that alternate between close and further away, and "are thought to use peripheral or extra foveal processes," which are somewhat unclear. Rockwell found significant differences between novices and experienced automobile drivers for the percent of time spent sampling lane position cues -- the experienced drivers spent a small, constant amount of time sampling such cues, and novices sampled far longer and more randomly, depending on speed. Rockwell
suggests novices don't make good use of peripheral processes, but they can learn to do so.

**Information Processing: Adopting a Paradigm**

A "paradigm" is an often misunderstood construct brought into the science of the last two decades. While the concept's popularity is attributed to Kuhn (1962), who has written extensively about paradigms, others point out that it is not at all clear exactly what is meant by paradigm. Masterman (1974) shows that Kuhn himself uses "paradigm" 21 different ways, in three general categories. At its simplest, a paradigm is a "way of seeing", an organizing principle that supplies tools to solve scientific problems. A paradigm is a "framework for discussion" (Popper, 1974). While most contemporary authors agree that a scientific community shares a paradigm, it appears that current motor vehicle research lacks this unifying framework. Nevertheless, it does seem that adoption of a paradigm -- an organizing principle -- might give assistance in structuring principle data on motorcycles -- rational interpretation of experimental results.

One paradigm popular with psychologists and human factors engineers is "information processing", and this
paradigm holds promise for interpreting findings in this research. However, placing the motorcyclist's perceptual tasks into an information processing paradigm requires operationalizing certain terms and concepts, and requires a bit of history review, too.

The development of information theory is attributed to Shannon and Weaver (1949), though some concepts such as "cybernetics" can be traced to Weiner (1948). Most developments in information processing derived from attempts in the second world war to improve transmission speed and accuracy of all manner of man-machine communications. (The birth of the computer was part of this intensive wartime activity.) DeGreene (1970) and others suggest that the wartime needs of "operational deficiencies in bombing, artillery targeting, submarine sonar detection and training, lead psychologists, mechanical and electrical engineers, linguists and others to join forces on typically human questions." (p. 119) In the last four decades, "human factors engineering" applies the basic information obtained by engineering psychology, applied physiology and related sciences to practical problems" and now includes aspects of visual perception (Forbes, 1972). As it has developed, the information-processing paradigm offers a structure for interpreting human factors data about motorcycle operators, and many provide useful answers to the
question, "Can man and machine co-exist safely?"

According to McCormick (1970), information is a reduction of uncertainty, or more simply, the unambiguous answer to a question. Engineers and psychologists measure information in bits, the "amount of information we obtain when one of two equally likely alternatives is selected" (McCormick, 1970).

Channel capacity refers to the limits of information that can be received through one or more sensory modes and processed for importance within a given time. Broadbent (1958) says that with multiple sensory inputs, there is a marked loss of information due to channel capacity. In other words, a person's channel capacity is a bottleneck that eliminates some information the same way a baseball player could only catch a certain number of simultaneously struck pop-flies. Channel capacity is the upper limit of information an individual may process at one time.

McCormick (1970) states his position on channel capacity quite succinctly:
"Let's face it -- we humans do not operate at perfect efficiency. There are certain kinds of situational variables that have some bearing upon one's performance, some for the worse (for example, noise), others for the better. Thus, it might be useful to pin down the situational elements that are related to information processing-handling abilities of people. If we could do this, we might, in the design of systems, capitalize on those features or schemes which facilitate information handling."

McCormick lists ten variables that affect the visual and auditory human information channels:

* Alphanumeric characters
* Colors
* Visual angle
* Size of form
* Brightness of lights
* Flash rate of lights
* Sound frequency
* Sound intensity
* Sound duration
* Sound direction

A motor vehicle operator has to 1) scan and identify combinations of the above signals to determine which carry useful information, and reject those that do not; 2) choose which signals are threatening to his safety and which are
not hostile; 3) execute an action based on that discrimination; 4) do it all within a short reaction time. Of course, the operator is equipped with a brain capable of hundreds of minute "decisions" per second, but applying Broadbent's concept of channel capacity, there are still upper limits to how much input can be processed at a time by an individual. A novice motorcycle operator, for example, is operating an unfamiliar vehicle in a threatening environment. The operator's channel capabilities may be exceeded and consequently the driver loses the ability to process and act on all information at the necessary speed. The result may be an inappropriate response, or no response to threatening stimuli. The operator may be scanning the environment rapidly but inefficiently, as Rockwell's research shows. McCormick's model seems useful in interpreting what a motorcyclist does in traffic when faced with a multitude of stimuli.

Another important element of human information processing is the concept of time-sharing, in which an individual has more than one task at hand, and has to attend to all of them simultaneously. Humans constantly shift focuses from one task to another rather than attend to all of them simultaneously. There is apparently no simultaneous doubling up. In motor vehicle operation, time-sharing means switching from one task to another, such as monitoring the visual field,
then monitoring the auditory field, then monitoring the machine. Switching must be quick and it must be accurate. Theoretically, an operator could be taught a hierarchy of important and less-important elements to attend to and thus improve the efficiency of time-sharing and his safety.

Olson (1959) examined the nature of time-sharing in a study wherein the subject simultaneously guided a vehicle on a mock roadway, while paying attention to identifying six to eighteen pointers on dials which moved randomly. Olson shows that performance on both the tasks was correlated very poorly ($r = .20$). As McCormick (1970) relates about the Olson study, "When the pressures of speed-stress tax the capacities of people [in a time-sharing task], something has to give, specifically performance on some time-shared tasks." He continues, "It is fairly evident that there are bounds beyond which the time-sharing of sensory inputs typically result in some degradation of performance" (p. 121). It is into this context that the present study is cast: experienced and trained motorcycle operators should show more adaptive strategies and more efficient time-sharing in a simultaneous task. They should be able to accurately sample the visual field without great difficulty, time-sharing with the task of vehicle guidance. Inexperienced and untrained operators might exceed their capacities to input, process, and act on novel
stimuli. The result may be rapid but inefficient scanning and slow reaction times; these are apparently the problems in the experimental literature on motorcycle accidents.

Alexander and Lunenfield (1975) made an interesting contribution to the information processing paradigm of motor vehicle operation. They report that the time-shared tasks in an auto driving situation can indeed be arranged hierarchically. Steering and speed control are at the highest, most important level, making responses to road situations or traffic at the middle level and route or trip planning preparation are at the lowest level in the hierarchy. In terms similar to "time-sharing", these authors suggest drivers "load shed", that is, focus their attention on higher hierarchical levels as a particular situation demands. They eliminate competing tasks and pay attention to the points at hand. Threatening information would dominate information of a more casual character in an operator's responses.

Ideally, any educational experiences under controlled classroom conditions should adapt, or transfer, to new situations. When a practiced response transfers to a new situation, transfer is said to be positive. Alternatively, training is useless if a new response is made to an old stimulus; transfer is said to be negative in such instances. Simulators
are good examples applying the principles of transfer of training. A driver reacts to realistic stimuli under controlled situations, and the reactions should transfer to the real environment if transfer is positive. McCormick's transfer surface illustrates the relationship between response similarity and amount and direction of training transfer. See Figure 5.

Figure 5.
As McCormick notes of the Osgood illustration, "The amount of positive transfer is shown to be optimum when both the stimulus and the response of the transfer task are identical to those of the training task." This would be the case in motorcycle training if the educational activities were conducted under real life conditions, complete with actual intersections and potentially threatening traffic. Since this is not the case and driver education takes place in a classroom with or on a well controlled driving range, transfer of training to real life traffic is necessarily less than 100%. In really poor educational settings, such as learning to drive in a field or an alley, transfer may range further downward. With no training at all, there is no chance to learn to adapt to real traffic by practicing under safe conditions. There is no transfer with no training.

A real-life operator drives in a complex environment. A schematic description is described by Rockwell (1972) and is reproduced here as a rather mechanical system, operator controlled and flexible. Figure 6 illustrates how very complicated and interactive driving is. To be successful, a driver education curriculum must utilize much of this model including stress and proprioceptive cues, and allow that training to transfer to real life. Using the information-processing paradigm researchers can quantify these cues, tie them to
concepts of transfer of training and channel capacity, for example, and rationally interpret empirical observations.
Figure 6.
Kidd, (1962, in Gagne) summarizes typical human errors in equipment operation and also suggests their possible causal factors. Table 4 illustrates these categories, and especially important to this discussion are: 1) errors in signal detection (too many significant signals); 2) error in action selection (matching of actual and required patterns faulty; and 3) errors of commission (action-control relationship not understood by operator).

Table 4

<table>
<thead>
<tr>
<th>Typical Human Errors in Equipment Operation and Their Possible Causal Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failure to detect signal</td>
</tr>
<tr>
<td>Input overload</td>
</tr>
<tr>
<td>a. Too many significant signals</td>
</tr>
<tr>
<td>b. Too many separate input channels</td>
</tr>
<tr>
<td>Input underload</td>
</tr>
<tr>
<td>a. Too little variety of signals</td>
</tr>
<tr>
<td>b. Too few signals</td>
</tr>
<tr>
<td>Adverse noise conditions</td>
</tr>
<tr>
<td>a. Poor contrast</td>
</tr>
</tbody>
</table>
b. High intensity of distraction stimuli

Incorrect identification of Signals
- Code form or typology unclear
- Lack of different cues
- Inappropriate filtering set (expectation)
- Conflicting cues
- Conflicting identification requirements

Incorrect value-weighing or priority assignment
- Nonlinear predictions required
- Multiple of complex value-scaling required
- Values poorly defined or understood
- Contingencies poorly defined

Error in action selection
- Matching of actual and required patterns faulty
- Consequence of courses of action not understood
- Appropriate action not available
- Correction action inhibited
  - a. Cost considerations
b. Procedural Prohibitions

Error of commission

Correct tool or control not available
Action-control relationship not understood by operator
Action feedback unavailable or delayed

(After Kidd, in Gagne, 1962)

Static visual acuity is described by Burg (1964) as the ability of an observer to discriminate an object when there is no relative movement between the observer and the object itself. Dynamic visual acuity, on the other hand, refers to a situation with relative movement -- either the object or the viewer is moving. Burg (1964) suggests there may be little correlation between dynamic and static visual acuity of automobile drivers, but did not test motorcycle operators under either condition. Burg and Hulburt (1959, p. 196) say that, "A person's ability to discriminate a moving target cannot be predicted as adequately from his static acuity, and the adequacy of this prediction decreases as the speed of the moving target increases." The application of this finding is especially important to motorcycle users, where novices with little or no screening for ability (and perhaps only testing for static acuity in a state-operated driver's license facility) are in control of a rapidly moving motor
vehicle with visual demands in all directions. Novice riders may not be able to fulfill any of the required task demands, and routine static or dynamic visual testing might reveal these inadequacies.

**Developing Visual Skills in Drivers**

Unfortunately, there is little literature available on applying experimental principles to improving visual skills in motor vehicle drivers, and almost none about motorcycle operators. Mourant and Rockwell (1970) examined beginning high school drivers and found that at least some visual skills developed quickly during the training. These included velocity control and car-following; the authors showed that eyesearch and scan capability continued to develop even after the training ended.

McPherson (1966) showed a series of 35mm slides depicting various traffic situations for a short amount of time. Driver training students were able to increase their ability when the experimenter artificially decreased the "response-window time". The author concludes that students could learn to take in more visual information under the experimental conditions. Extending this finding, it is possible that motorcycle operators, too, can learn to take in more visual
information visually.

Rockwell (1969) developed a system for improving visual information input to drivers. His "head-up display" was fixed in an existing automobile to provide route-guidance information. The author does not say how successful his system was, but that the method of focusing a visual display on the windshield at an infinite focal distance might have enabled the driver to pay attention to the task and take in more visual information at the same time.

Motorcyclist Training in the United States and Canada

Standard courses of instructions for automobile drivers have been available since Professor Amos Neyhart developed the first "behind-the-wheel" course at Pennsylvania State University. Today, almost every high school in the country teaches driver education courses, as do private companies and even some larger department stores.

Motorcyclist education programs have been much slower to develop. In 1973, a half decade after Neyhart, the five major motorcycle manufacturers pooled their resources and formed the Motorcycle Safety Foundation (MSF), whose goals were to reduce traffic injuries and fatalities through training. According to the former president of the MSF, the Motorcycle
Safety Foundation "was developed by educators and experienced motorcyclists and represents the best thinking and practice in motorcycle rider education at the time." (Hartman, 1976).

At this time, there is little doubt about the preeminent position of the MSF in standardized motorcyclist training and curricula. Every state in the country has at least one operational site, while California has nearly 50 such sites. Ohio has been slow to develop MSF training sites, even though it is the fifth largest motorcycle registration state in the nation with over 300,000 registrations (AMA, 1984). There are only two currently operating sites in Ohio. Ohio trained only about 200 students in 1984, whereas Illinois trained over 4,800, with an almost identical motorcycle registration population of 295,000.

The MSF's standardized Motorcycle Rider Course (MRC) is 12 hours in length and includes six, 16-millimeter films, textbooks, instructor's guides and evaluation forms for teachers. The MSF trained about 42,000 students in 1984, and has trained about 400,000 civilian and military students in its 10-year existence.

In Canada, the Canadian Safety Council has developed a similar course called the Motorcycle Training Course (MTC). According to the MTC coordinator, there is little difference between the MRC and the MTC, although the MTC spends more
time practicing on the motorcycles. The MSF spends about half the time in the MRC on on-cycle training, compared to the MTC's 85 percent, according to Fassnacht, (personal communication, May 24, 1984).

Summary and Critique of Relevant Literature

In analyzing the experimental literature associated with motorcycle safety, perception and accident avoidance, four distinct categories for discussion emerge. These categories are not mutually exclusive, but fall logically together.

First, almost no basic research in basic perception and motorcycle accident avoidance has yet been accomplished. Much more literature exists on the human factors associated with auto driving, but it is risky to generalize from one population to another, especially when the task demands, accident-descriptors and user-demographics are so different.

With the need for improved motorcycle safety, it is easy to argue that almost any scientifically sound experiment in motorcyclists' human factors could potentially have lifesaving benefits. For this reason alone, there is justification for basic research in motorcycle operator human factors. Since driver research began first with static displays and proceeded to dynamic displays, once the parameters of static
displays had been established, motorcyclist research should logically begin there, too, to establish similar research parameters. Dynamic human factors research in motorcyclist perception can follow.

Second, traffic safety educators often refer to the "benefits" of education in motorcycle operation, but experiments or studies have not been generated which unequivocally support the effectiveness of rider training. Legislative bodies and enthusiasts alike sing the praises of education, but why? The literature shows there is even evidence that motorcycle education has negative effects. Human factors research strongly supports the importance of visual information acquisition, speedy decision-making, yet training curriculum writers are not in a position to say exactly what elements of their motorcycle education program are beneficial, which could be improved, and which could be reduced.

Third, the literature shows many laboratory findings based on visual discrimination tasks made with no secondary task employed. For example, a visual task is assigned to subjects without regard to time-shared motor tasks demands such as guidance (steering). It may be that single channel tasks (visual only, for example) are probably not cognitively taxing as in real life. The research of Burg (1964, 1966), Rockwell (1969, 1972), and McPherson (1966), for example,
examine visual capacities but do not require the subject to drive a simulated auto, or even apply a mock brake intermittently while their information-processing capacities were investigated. The primary visual task may not be difficult without a secondary task. Under circumstances of competing tasks, subjects may have to learn to time-share, and under the stress of a secondary motor task, they may show degraded performance on either or both. It can be argued that findings in past literature are artifactual and would not replicate in a real world setting because there was no secondary task employed.

Fourth, the automotive human factors literature (and motorcycle literature to an even greater degree), does not seem to have a unifying theme to interpret and organize findings. Some ideas which seek to unify, clarify, and guide curriculum development like the "IPDE process" are dead-end streets, and can't be directly applied to real-life settings any more than can the theory that operant condition is responsible for driver responses. They don't explain phenomena, but only describe them. While such acronyms are intuitively appealing, they offer no useful explanation for observed results, no model of human performance, and no suggestions for further research.
(IPDE is an acronym for Identify, Predict, Decide, and Execute (Automotive Safety Foundation, 1972)).

Information processing provides a central theme, or a paradigm, to frame results in, create models and offer explanations. With good basic empirical data on motorcyclists' perceptual capabilities, an information-processing paradigm can be applied to motorcycle research. As Atteneave points out, "information theory is not going to provide a ready-made solution to all psychological problems, but employed with intelligence, flexibility, and critical insight, information theory can have great value in the formulation of certain psychological problems, and in the analysis of certain psychological data" (Atteneave, 1970).

There are, of course, mild dangers in posing questions in terms of channel capacity, uncertainty, and transfer of training especially in the early stages of research, because the phenomenon observed in basic research may not replicate. Yet there seems to be a greater danger in letting observed phenomena float unattached, or in structuring those phenomenon in an arbitrary framework that has no inferential character.

To summarize, the literature reviewed here strongly suggest a study of whether educating motorcyclists makes a significant difference in their ability to make accurate responses, and make them quickly. In more operant terms, the
present study will examine the static visual acuity of motorcyclists who vary in levels of training, and will require them to time-share the visual task with a motor task simultaneously. Trained motorcyclists should be better able to "time share" both tasks.

If the results of the study support the implicit hypothesis that trained motorcyclists make faster and more accurate responses, then the study will provide an empirical rationale for expanding training opportunities.
Chapter III

RESEARCH PROCEDURES AND METHODS

The purpose of this chapter is to describe and justify the research procedures and methods used to collect data on the problem of the visual acuity motor skills and reaction times of trained and untrained motorcyclists. The chapter includes sections on the use of simulation, descriptions of all apparatus and stimuli, the research design, the use of control groups, experimental procedures and instructions to subjects. A description of the pilot study is also presented, and an analysis of its findings.

The Uses of Simulation

"Simulation, models and games are interrelated abstractions, idealizations or analogies of the real world," according to DeGreene (1970). Airline pilot trainees and millions of automobile novices are trained on simulators to create real-world analogy without endangering the trainees with hazards encountered in flight or in driving. Simulators allow a careful laboratory analysis of selected human factors variables under controlled conditions.
Advantages of motor vehicle simulation are that they don't subject trainees to danger, and provide a less costly substitute for real life operations, such as in space flight simulation. However, sometimes simulation offers a "low fidelity" to the actual machine; it is a simulator only at face value, and this is a distinct disadvantage. Perhaps, for example, a flight simulator cannot duplicate weightlessness, which would in turn affect the way a trainee functioned.

While automobile simulators have been in use for decades, there are no regularly used motorcycle simulators, according to Thom (personal communication, May 14, 1984). As a result, there is no readily available research upon which to base the construction of the simulator used in this research.

However, as McCormick suggests, simulation is still useful under certain conditions. These conditions include tasks that can be identified and "spelled out to the trainee, based on a task analysis, and tasks which provide psychologocal simulation or task learning" (McCormick, 1970, p.229). The motorcycle simulator in this research meets these criteria.

In this project, tasks are spelled out clearly on a tape recorded set of instructions (See Chapter III: Instructions to Subjects). Motorcycle simulation tasks used in this research (clutch, brake, horn, and throttle) are identified elsewhere in a task analysis by McKnight and Heywood (1974),
and the motorcycle simulator is entirely oriented to transferring training to real life. The research described here uses a partial-motorcycle simulator in that only a handlebar is used; the simulator does not provide standard foot controls. However, since the handlebar apparatus meets McCormick's criteria, it is assumed that the simulator has at least moderate fidelity, and thus is justified for research.

SUBJECTS

This study compares the visual and motor skills and reaction times of novice and trained motorcyclists to provide an empirical rationale and a basic human factors interpretation for motorcyclist training. It is useful and economical to expand this two group design by adding experienced motorcyclists on one end of the continuum, and non-motorcyclists as a control group on the other end. These additions broaden the available interpretation. The four categories of subjects used in this study are:

* ER- Experienced Motorcycle Riders
* FTN- Formally Trained Novice Motorcyclists
* NOV- Novice Motorcyclists
* NON- Non-Motorcyclists
Subjects can not be randomly assigned to groups because they have "self-selected" themselves according to the above classifications. Samples of the ER, NOV and NON populations were randomly chosen from these strata, and all were volunteers. Riders in the FTN strata above were chosen at the April 13-27, 1985 training site at Cuyahoga Community College; this was the single operating motorcyclist training site in Northern Ohio.

Subjects in groups labeled NOV, FTN, and NON were not required to have a motorcycle license but had to meet the requirements, since they were potential licensees. Non-motorcyclists were not required to have any license, but were also considered potential motorcycle riders. Experienced riders, of course, had to have a current motorcycle license.

Sixteen pilot subjects were used in this research; there were 4 pilot subjects in each of the 4 test categories, and all were randomly selected from the population of Columbus, Ohio motorcyclists, except for the FTN pilot subjects which were selected from the population of all students being trained at Cuyahoga Community College located in Cleveland, Ohio. Pilot subjects were tested during the Motorcycle Rider Course (MRC) held April 13-27, 1985 at Cuyahoga Community College.
Forty test subjects were used in this full project. There were 10 subjects in each of the four test categories, and all were volunteers selected from the population of Columbus, Ohio motorcyclists, including the 10 FTN subjects who were selected from the population of all students being trained at the Cuyahoga Community College located in Cleveland, Ohio. The FTN subjects consist of all the students undergoing voluntary training at the Cuyahoga Community College; the training consisted of the 20 hour Motorcycle Rider Course.

**Visual Task Apparatus**

It is important to have reliable apparatus that lends itself to experimental replication. In this experiment, the apparatus needed to present the stimuli to all subjects consists of:

* a Kodak Carousel slide projector
* a Lafayette Instruments electronic shutter timer
* a Lafayette Instruments lens shutter.

Slide presentations were uniform at 50 milliseconds, as this exposure is well supported by the experimental literature. Slide exposures were controlled by a slide projector
lens-mounted shutter, which was in turn controlled by a timer. Slides were manually advanced. All visual targets were viewed on a flat white non-reflective paper at a distance of 12 feet. A Commodore VIC 20 controls all operations, and initiates each slide at the command of the experimenter. A computer program (listed in Appendix A) controlled all exposures and displayed the motor response type and reaction time on a cathode ray tube.

Motor Task Apparatus

Four independent, simulated motorcycle handlebar tasks, prepared for this study were:

* a motorcycle brake lever
* a motorcycle clutch lever
* a motorcycle horn button
* a motorcycle throttle.

All of the simulator's controls were connected to a 5 volt electronic interface (diagrammed in Appendix B) which, in turn, connected them to computerized timers. A Commodore VIC 20 was programmed to display on a cathode ray tube the
control manipulated by a subject, and also displayed the reaction time for that control. Slides were manually advanced. A schematic of the experimental apparatus appears in Appendix C.

**Stimuli**

A stimulus can be defined as "something that excites an organism to functional activity" (Urdang, 1973, p. 1291). A response to a particular stimulus means "any behavior of a living organism that results from stimulation" (Urdang, 1973, p. 1125). In studying human behavior, stimuli used have ranged from electric shock to food, to psychological punishment. Typical responses in behavioral studies are galvanic skin responses, electroencephalograph readings, or verbalizing a learned word pattern.

In the study of human memory, cognitive psychologists have commonly used arrays of English-alphabet letters, and Arabic numerals as stimuli. George Sperling, one of the pioneers in human cognition, used visual arrays of letters, exposing them briefly in rows and columns for about 50 milliseconds (.050 sec.). (Sperling, 1967)

Other types of stimuli are commonly used in cognition studies such as lighted rectangles which differed by length and brightness, or words flashed at very short (.050 sec.)
durations (Egeth, 1967). Neisser used words embedded in long lists (Neisser, 1963). In other studies, he used embedded letters and numbers (Neisser, Novick, and Lazar, 1963) exposed at brief durations.

Another type of stimulus, called a Landoldt ring, has been used. Created in 1906 to replace the Snellen eye chart for illiterates, Landoldt used a specially formed letter "C" with a gap orientation that varied up, or down, or left or right. These stimuli have been used in various experiments to test the efficiency of hemiretinal of viewing (Markowitz and Weitzman, 1968), and laterality differences in humans (Winn, 1975).

Stimuli used in the present study were similar to visual stimuli used by Sperling, Neisser, Markowitz and Weitzman, and others in that they are alphabetic (English letters), numeric (Arabic numerals), or spatial (Landoldt rings) stimuli, all presented at short intervals, typically .050 seconds. In this research, the one notable addition to the stimuli used in the experimental literature was the use of arrow stimuli, which, like the Landoldt ring, were oriented pointing up, down, left, or right.

For this study, stimuli were Letraset-brand, press-on black characters, all in Helvetica medium, 30 point type. The stimuli were photographed on a white background which was
lighted with studio lights. Photography was done with a Nikormat camera, at 1.8 feet. A standard studio gray card was used to determine aperture setting (f=2.8); exposure duration was .250 seconds. Kodak tungsten film was used throughout, ASA 160. Viewed on the screen as slides, each stimuli had high contrast, and appeared about 1 1/2" high when viewed at 12 feet.

There were 16 different stimuli used: there were 4 letter stimuli (B, L, M, and T) chosen for low confusability (see Kinney, Marsetta and Showman, 1966). There were also four numerals (1, 3, 5, 7) chosen for low confusability. There were four arrows positioned up, down, left, or right, and there were four Landoldt rings whose gaps were oriented up, down, left, or right. (Examples of these stimuli appear in the Appendix.)

Visual stimuli were arbitrarily linked to motor tasks. Upon presentation of a letter stimulus, the subject was required to name the letter, and simultaneously pull in the clutch lever. When the subject was presented with a number, the subject named it, and pushed the horn button. If the subject was presented with an arrow, the subject named the direction and squeezed the brake lever. Finally, when the subject saw a Landoldt ring, the subject named the gap direction, and twisted the throttle about half a turn. These links never
varied during the study. The purpose of the arbitrary linkage was to ensure that no subject group had a pre-test skill advantage, especially the ER group; that is to say, all subject groups were required to learn a new task. Groups only differed on pre-test training, which was the independent variable. To ensure that subjects have understood the instructions, each was allowed 4 trials for practice. During practice, subjects were first shown a letter stimulus, and asked to make the appropriate motor response, which was a clutch lever pull. Then, each subject was shown one numeral, then one arrow and one Landolt ring and each time asked to name the stimulus while making the appropriate motor response. Each of the 16 test stimuli were presented in the central visual field, but the stimuli were also shown 7 degrees to the right, and 7 degrees to the left of the central point. Thus, a total of 48 stimuli were used, and their order were randomized.

Research Design

A true experiment requires administering an independent variable to groups that are randomly selected to receive it, and a control group which is not. Campbell and Stanley (1972), however, point out that
'There are many natural social settings in which the research person can introduce something like [true] experimental design into his scheduling of data collection procedures (e.g., the when and to whom of measurement), even though he lacks full control over the scheduling of experimental stimuli which makes a true experiment possible. Collectively, such situations [which lack full experimental control] can be regarded as quasi-experimental designs.'

One of Campbell and Stanley's quasi-experimental designs is called the "Nonequivalent Control Group Design (Campbell and Stanley, 1972, p. 47). In this design, the groups constitute naturally assembled collectives, such as classrooms, as similar as availability permits, but the authors suggest retaining the pre-test if possible. "The assignment of [the independent variable] to one group or the other is assumed to be random and under the experimenter's control," according to Campbell and Stanley.

The design of this research is quasi-experimental because all conditions required by a true experiment, and in particular, random subject selection, can not be met. This is because subjects in the population "self-select" themselves to be motorcyclists, or not, and to seek training. This research, then, is a Non-equivalent Control Group design as described by Campell and Stanley with the exception of a pre-test, since a pretest was not possible. The experimenter did not know who would be in the FTN group until the date of the course. Beyond this, all standard experimental procedures
apply even to quasi-experimental research.

While a true experiment is more able to show cause and effect relationships between dependent and independent variables, quasi-experiments have varying degrees of confirmation [of theories] depending on "the number of plausible rival hypotheses available to account for the data" (Campbell and Stanley, 1972, p. 36). "Where controls are lacking in a quasi-experiment, one must, in interpreting the results, consider in detail the likelihood of uncontrolled factors accounting for the results. The more implausible this becomes, the more "valid" the experiment" (Campbell and Stanley, 1972, p. 36). Because this research is a quasi-experiment, rival hypotheses must be addressed as they may account for the observed results.

**Variables**

The **independent** variable was "Training" in four levels; ER (Experienced Rider), FTN (Formally Trained Novice), NOV (Novice), and NM (Non-Motorcyclist). The **dependent** variables were: 1) Errors on the Motor Task, 2) Errors on the Visual Task and 3) Reaction Time measured on the Motor Task.
Control Groups

The need for a control group is well documented in the literature (D'Amato, 1970; Underwood, 1957; Boring, 1954). The control group is typically the group not administered the independent variable or is composed of "subjects serving under zero amount of the independent variable, or under a value that is in some sense a standard value" (D'Amato, 1970). In this research, the control group was the NON group (Non-motorcyclists). This group's subjects had a zero level of the independent variable training. Subjects had self-selected themselves into groups quite independent of this research—that is, they had chosen to be motorcyclists or not. Those that had chosen not to be a motorcyclist could be sampled in the Non-Motorcyclist strata.

Test Procedures and Instructions

Subjects were solicited by word of mouth, and by invitation. Scheduling for testing is made a week in advance. Upon arrival, each subject was greeted, and seated behind the simulator. Then the subject was asked to listen to the tape recorded instructions while his or her eyes adjusted to low lighting conditions. All subjects were questioned about their
motorcycle experience, and that was noted.

Instructions were as follows:

Hello. Thank you for offering to assist us in our research on driver reaction times. Today, you will be shown a series of slides that will be flashed on the screen, and you will be asked to identify them. At the same time, you will be asked to operate the apparatus in front of you. The task is simple, but will probably require some practice trials before you feel comfortable in starting. You should focus on the center of the screen, where there will be four kinds of slides flashed: there will be letters, (pause) numbers, (pause) arrows, (pause) and rings. (pause) Look at the card supplied (wait 5 seconds). The letters and numbers are identified by simply calling their names. On the arrow and the rings you identify which way the arrow is pointing, or which way the ring is open. On the card, the correct responses would be T, 5, up, and right. Of course, there will be only one letter or arrow on a slide. Do you have any questions so far? (wait 5 seconds).

Now we're almost ready to start. Each time you see a letter, you should name it and simultaneously pull in the clutch lever. It's on the left handlebar. Try it. Just push it all the way in. That's for letters. Whenever you see an arrow on the screen, you should call out its name and pull the brake lever on the right handlebar. Try it. That's for arrows.

Whenever you see a number on the screen, you should call out its name and push the red horn button located on the left bar. Try the horn. That's for numbers.

Finally, when you see the ring, name its direction and twist the throttle about halfway back, located on the right handlebar. Try it once. That's for rings.

OK. Let's try some practice trials until you get the feel of the test situation. This may sound complicated now, but it will get much easier with practice—everyone else says so, anyway.
If you're not sure, you must guess, but it's better to take a bit longer to make sure. Try to be as accurate on the slides as you can.

OK. Let's proceed to the practice trials.
(4 practice slides)
(Assume any questions)
(Begin 48 test slides)

Each subject was given four unscored practice trials, and then given an opportunity for final questions to be answered. Next, the test set of 48 slides were shown, and 2 responses were noted for each trial. A total of 48 stimulus slides were shown, and a pair of simultaneous responses (one identifying the visual stimulus, and one that was linked to a motor response on the simulator) were recorded. A sample score sheet appears in Appendix E.

Each subject signed an "informed consent" form numbered HS-027 by the Ohio State University (see Appendix F).

The Pilot Study

A pilot study is a complete preliminary research project undertaken to determine parameters of independent and dependant variables, and to smooth out all procedures. Among the important questions asked during the pilot study were:

* are the tasks too demanding of subjects?
* Are four practice trials enough to familiarize the subjects with tasks?
* Are the instructions clear and concise?
* Does the apparatus work properly?
* Can reliable visual task scores be taken?
* Can reliable motor task scores be taken?
* Can reliable reaction times be taken?
* Is the visual stimulus exposure time too fast?
* Could oral debriefing of pilot subjects yield clues to improve procedures, methods or apparatus failures, or reveal fatigue?

During the initial shakedown of the pilot study, a group of 8 experienced riders was chosen and tested to give extensive debriefing and comments. During this first phase of the pilot study, procedures were lax, scores were not recorded, and adjustments could be made at will even during the test.

In the next phase of the pilot study, 4 subjects in each group were tested under true, controlled test procedures. No formal data analysis was made, but a check was made to determine whether the data conform to the hypothesis. If further procedural adjustments were not necessary, the full study, with 40 subjects, would be undertaken.
Pilot Study Results

During the first phase of the pilot study, it was determined that ambient lighting would have to be kept to almost total darkness, but a small lamp was permitted to allow safe movements in the laboratory. The cathode ray tube was causing some distraction to early pilot subjects, so it was turned away from them. Some adjustments were made to the simulator to clutch tensioning springs to match the "on" point of the brake lever. Beyond these adjustments, the following answers were found in the pilot study:

* tasks were not too demanding of subjects;
* four practice trials were enough to familiarize the subjects with the controls;
* the instructions were clear and concise;
* the apparatus worked properly and reliably;
* reliable reaction times could be measured;
* reliable error scores could be taken;
* visual stimulus exposure times were not too fast.
* a short oral debriefing of pilot subjects revealed that they were not fatigued, and that the procedures weren't overly taxing. Debriefing only revealed that the tasks were not too difficult or confusing, even for Non-Motorcyclists.
Average reaction times were fastest for the Experienced Riders, next fastest for Formally Trained Motorcyclists, next fastest for Novices, and slowest for Non-Motorcyclists. The pilot study revealed that error scores on the Visual and Motor tasks were best for Experienced Riders, next best for Formally Trained Motorcyclists, next best for Novices, and worst for Non-Motorcyclists. Because these pilot data were in the direction of the hypothesis suggesting differences among groups of riders on the factor of training, it was determined that a full study could be undertaken. Figure 7 shows that the implicit hypothesis is supported, and that a very small sample of motorcyclists do vary on the factor of training.
Summary of Research Methods and Procedures

The objective of this research is to determine whether there are reliable differences in visual and motor skills and
reaction times between Novice and Formally Trained Novice motorcyclists. Experienced Riders and Non-Motorcyclists are added to extend the data findings in either direction, and to gain an economical control group in the case of the NM group.

Because it was not possible to randomly assign subjects to levels of training, the independent variable, the design of the current research is necessarily "quasi-experimental." The true control group was composed of novice motorcyclists (NOV), who had a zero level of the independent variable. The literature supported the use of a simulator to test whether a training in the form of motorcycle education course affects errors on a Visual Task, a Motor Task, and Reaction Time. Stimuli used in this research were also supported by past studies.

A pilot study using a small number of test subjects was undertaken, and a cursory analysis of the pilot data revealed that the hypothesis was indeed supported. As a result, a full study was undertaken, and discussed.
Chapter IV

RESEARCH RESULTS

The objective of this research was to establish a link between motorcycle training and an improvement in visual skills, motor skills, and reaction times. Yet, as Babbie (1973, p. 47) says, "It is normally impossible to arrive at a wholly unambiguous and completely acceptable measure of any variable." In addition, Campbell and Stanley have pointed out that in a quasi-experimental study, competing hypotheses must be identified and addressed. The fewer rival hypotheses there are, the more likely the observed results are due to the independent variable (treatment).

To provide the least ambiguous answers to a given question, researchers establish a hierarchy of statistical analyses, beginning with whether the research is reasonably valid and reliable. Validity refers to "the degree to which the test is capable of achieving the aims or purposes it is intended to serve" (Tiffin and McCormick, 1965, p. 127). "Reliability, which is necessary in a test because it limits the validity", refers to "the degree to which a test measures consistently whatever it does measure" (Tiffin and McCormick, 1965, p. 129). An item analysis was performed, and this was followed by a test to establish whether subjects learned the
task adequately. A comparison was made between the first ten and last ten items on all tasks. Next, a test of the assumptions of the analysis of variance was conducted. Finally, a multivariate analysis of variance was performed which established levels of significance for three dependent variables simultaneously, and then separately in univariate analyses of variance. Pairwise comparisons were made among the four groups on any of the three dependent variables that showed significance.

**Demographics of Test Subjects**

To examine the argument that preexisting differences in subject groups accounted for observed differences, an examination of these groups was made. Age variation, motorcycle operator experience and gender proportions were investigated, and these results are displayed below in Table 5.
Table 5
Demographics of Test Subjects

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean Age (In Years)</th>
<th>Motorcycle Experience (In Months)</th>
<th>Gender (Proportion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>34.6</td>
<td>178.8</td>
<td>.90 Male</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.10 Female</td>
</tr>
<tr>
<td>FTN</td>
<td>28.7</td>
<td>8.7</td>
<td>.50 Male</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.50 Female</td>
</tr>
<tr>
<td>NOV</td>
<td>33.8</td>
<td>6.4</td>
<td>.20 Male</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.80 Female</td>
</tr>
<tr>
<td>NM</td>
<td>37.9</td>
<td>0.0</td>
<td>.20 Male</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.80 Female</td>
</tr>
</tbody>
</table>

Relative to age, there were no significant differences between the groups (p=.0923); most subjects were over 30, except in the FTN group. There were differences in proportions of males and females, however, with the ER group being almost all male, and the NM and the NOV groups being mostly female. These differences are discussed later, as possible sources of between group variance.

Groups differed greatly on pretest experience. As would be expected the ER group had been riding longest, with an average of 178.8 months; the FTN and NOV groups both had less than nine months of riding. The NM group had 0 months riding experience.
Validity Analyses

The first phase of data analysis is establishing face validity and internal validity. To establish face-validity for a test, it must appear valid. Because a motorcycle simulator is used to test mostly motorcyclists in this experiment, the task appears face-valid. However, it could be argued that the stimuli (alphabetic and numeric characters, Landolt rings and arrows) were not face-valid. They do not represent real traffic stimuli. Using a brake lever might be associated unambiguously with a photo stimulus of a child running into the roadway, but what unambiguous stimulus could be associated in a static display with the use of the clutch or throttle? Pilot demonstrations failed to indicate unambiguous static visual stimuli. Second, it was decided that all subjects should have to learn the association task to support the idea that training alone allows quicker and more accurate responses. In this way, no subjects would be at an advantage due to experience. The static stimuli thus represented a way to achieve this goal and retain face-validity, even if they did not represent actual traffic stimuli.

Next, an item analysis was performed for internal validity. This procedure measured the extent to which the composite index related to (or predicted) responses to all the items included in the index itself (Babbie, 1973, p. 266).
In this research, one experimental trial varied little from another except on type of stimulus.

**Item Analysis**

An item analysis helps determine the relative difficulty of a particular item among all groups (Downie and Heath, 1974), and how well a particular item discriminates among the ER, FTN, NOV, and NM groups. In the first case, this analysis provided the proportion of all 40 subjects who responded correctly on the Visual and Motor Tasks.

In the second and third cases, an analysis was made of the proportion of correct (1) Visual, and (2) Motor responses by group. Since Reaction Time was a continuous variable, without "correct" or "incorrect" response, reaction times were not considered in the item analysis. Table 6 reveals the proportions of correct visual and motor test items overall and by subject group. Overall, the range of correct items extends from a low of .792 on item 8 to perfect scores on items 12, 15, 22, 32, 36, 41, and 44. According to this analysis, the combination of letter stimuli and clutch response were relatively easy items, with four of the seven perfect item scores. On the other hand, the combination of Landolt rings with
throttle items and arrows with brake items were the most difficult items with four and five of the lowest item scores, respectively. See Table 6.
### Table 6

**Item Analysis: Proportions of Visual and Motor Task Items**

**Correct by Subject Group**

<table>
<thead>
<tr>
<th>Item</th>
<th>Overall</th>
<th>ER</th>
<th>FTN</th>
<th>NOV</th>
<th>NM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.875</td>
<td>1.00</td>
<td>1.00</td>
<td>.80</td>
<td>.60</td>
</tr>
<tr>
<td>2</td>
<td>.979</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>.917</td>
<td>1.00</td>
<td>1.00</td>
<td>.80</td>
<td>.70</td>
</tr>
<tr>
<td>4</td>
<td>.854</td>
<td>.90</td>
<td>1.00</td>
<td>.80</td>
<td>.60</td>
</tr>
<tr>
<td>5</td>
<td>.979</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
</tr>
<tr>
<td>6</td>
<td>.979</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>.958</td>
<td>.90</td>
<td>.80</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>8</td>
<td>.792</td>
<td>.80</td>
<td>.80</td>
<td>.70</td>
<td>.70</td>
</tr>
<tr>
<td>9</td>
<td>.979</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>10</td>
<td>.833</td>
<td>.90</td>
<td>1.00</td>
<td>.60</td>
<td>.70</td>
</tr>
<tr>
<td>11</td>
<td>.958</td>
<td>1.00</td>
<td>1.00</td>
<td>.80</td>
<td>1.00</td>
</tr>
<tr>
<td>12</td>
<td>1.000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>13</td>
<td>.896</td>
<td>.80</td>
<td>1.00</td>
<td>.90</td>
<td>.80</td>
</tr>
<tr>
<td>14</td>
<td>.896</td>
<td>.90</td>
<td>.80</td>
<td>1.00</td>
<td>.80</td>
</tr>
<tr>
<td>15</td>
<td>1.000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>16</td>
<td>.917</td>
<td>1.00</td>
<td>1.00</td>
<td>.80</td>
<td>.80</td>
</tr>
<tr>
<td>17</td>
<td>.938</td>
<td>1.00</td>
<td>.90</td>
<td>.90</td>
<td>.90</td>
</tr>
<tr>
<td>18</td>
<td>.958</td>
<td>.90</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>19</td>
<td>.979</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>20</td>
<td>.896</td>
<td>.80</td>
<td>.90</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>21</td>
<td>.875</td>
<td>1.00</td>
<td>.70</td>
<td>.90</td>
<td>.80</td>
</tr>
<tr>
<td>22</td>
<td>1.000</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>23</td>
<td>.958</td>
<td>.90</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>24</td>
<td>.938</td>
<td>1.00</td>
<td>1.00</td>
<td>.70</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 7 compares the proportions of correct Visual Task items (1) overall and (2) by subject group. The proportion of correct visual stimuli ranges from a low of .938 on items 1, 8, and 10 to perfect 1.0 scores on 32 separate visual items. According to this analysis of the visual stimuli, Landolt rings or arrows, were more difficult for all groups to master. The distribution of incorrect items was too small to make inferences from (a sample of only three items that were .938), so a second analysis was done on any item’s visual error scores.

This distribution grew to a sample of sixteen items: 1, 3, 4, 8, 10, 11, 16, 17, 20, 21, 25, 27, 31, 37, 39, and 45. Of these, eight items were Landolt ring stimuli, indicating the relative difficulty of these test items. Four incorrect visual items were numerical stimuli, and three were arrow stimuli. Incorrect letter responses were not noted at all. On the whole, the Visual Task items were relatively easy for subjects to perform. See Table 7.
### Table 7

**Item Analysis: Proportions of Correct Visual Items by Subject Group**

<table>
<thead>
<tr>
<th>Item Overall</th>
<th>ER</th>
<th>FTN</th>
<th>NOV</th>
<th>NM</th>
<th>Item Overall</th>
<th>ER</th>
<th>FTN</th>
<th>NOV</th>
<th>NM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.938</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>25</td>
<td>.979</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
</tr>
<tr>
<td>2</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>26</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>.979</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>27</td>
<td>.979</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>4</td>
<td>.958</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>28</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>29</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>6</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>30</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>31</td>
<td>.958</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>8</td>
<td>.938</td>
<td>1.00</td>
<td>.90</td>
<td>.90</td>
<td>32</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>33</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>10</td>
<td>.938</td>
<td>1.00</td>
<td>.90</td>
<td>.90</td>
<td>34</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>11</td>
<td>.979</td>
<td>1.00</td>
<td>1.00</td>
<td>.90</td>
<td>35</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>12</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>36</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>13</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>37</td>
<td>.979</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>14</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>38</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>15</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>39</td>
<td>.979</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>16</td>
<td>.979</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
<td>40</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>17</td>
<td>.979</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
<td>41</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>18</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>42</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>19</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>43</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>20</td>
<td>.979</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
<td>44</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>21</td>
<td>.979</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
<td>45</td>
<td>.979</td>
<td>1.00</td>
<td>.90</td>
<td>1.00</td>
</tr>
<tr>
<td>22</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>46</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>23</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>47</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>24</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>48</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 8 reveals proportions of correct Motor Task items overall and by subject group. Overall, the range of correct items extends from a low of .854 on items 10, to perfect scores on items 12, 15, 19, 22, 32, 36, 38, 41, 43, and 44.

This analysis of Motor Task responses shows that clutch responses were easier items, with 5 of 9 subject score items. Among the difficult items on the Motor Task were three items below .90: items 1, 4, 8, 10, 13, 16, 20, 27, and 31. Among these, six items either were on the brake or the throttle. See Table 8.
### Table 8

**Item Analysis: Proportions of Correct Motor Task Items by Subject Group**

| Item | Overall | ER | FTN | NOV | NM | Item | Overall | ER | FTN | NOV | NM |
|------|---------|----|-----|-----|----|------|---------|----|-----|-----|----|----|
| 1    | .896    | 1.00 | 1.00 | .70 | .80 | 25   | .958    | 1.00 | 1.00 | .90 | .90 |
| 2    | .979    | 1.00 | 1.00 | .90 | 1.00| 26   | .979    | 1.00 | 1.00 | .90 | 1.00|
| 3    | .938    | 1.00 | 1.00 | .90 | .80| 27   | .875    | 1.00 | .80 | .80 | .80|
| 4    | .896    | .90  | 1.00 | .90 | .70| 28   | .958    | 1.00 | 1.00 | .90 | .90|
| 5    | .979    | 1.00 | 1.00 | 1.00 | .90| 29   | .958    | 1.00 | 1.00 | .90 | .90|
| 6    | .979    | 1.00 | 1.00 | .90 | 1.00| 30   | .938    | 1.00 | .90 | .90 | .90|
| 7    | .938    | .90  | .80  | 1.00 | 1.00| 31   | .896    | .70  | .90 | .90 | 1.00|
| 8    | .875    | .90  | .90  | .80 | .80| 32   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00|
| 9    | .958    | 1.00 | 1.00 | .90 | .90| 33   | .958    | 1.00 | 1.00 | .80 | 1.00|
| 10   | .854    | .80  | 1.00 | .70 | .80| 34   | .958    | 1.00 | 1.00 | .80 | 1.00|
| 11   | .979    | 1.00 | 1.00 | .90 | 1.00| 35   | .979    | 1.00 | 1.00 | .90 | 1.00|
| 12   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00| 36   | .000    | 1.00 | 1.00 | 1.00 | 1.00|
| 13   | .875    | .70  | 1.00 | .90 | .80| 37   | .979    | 1.00 | 1.00 | 1.00 | 1.00|
| 14   | .917    | .90  | .90  | 1.00 | .80| 38   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00|
| 15   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00| 39   | .979    | 1.00 | .90 | 1.00 | 1.00|
| 16   | .896    | .80  | 1.00 | .90 | .80| 40   | .979    | 1.00 | .90 | 1.00 | 1.00|
| 17   | .938    | 1.00 | .90  | .90 | .90| 41   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00|
| 18   | .958    | .90  | 1.00 | .90 | 1.00| 42   | .917    | 1.00 | .90 | .70 | 1.00|
| 19   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00| 43   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00|
| 20   | .896    | .80  | .90  | .90 | .90| 44   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00|
| 21   | .938    | 1.00 | .90  | .90 | .90| 45   | .917    | .90  | .90 | .90 | 1.00|
| 22   | 1.000   | 1.00 | 1.00 | 1.00 | 1.00| 46   | .958    | 1.00 | .90 | .90 | .90|
| 23   | .958    | .90  | 1.00 | .90 | 1.00| 47   | .979    | 1.00 | 1.00 | 1.00 | .90|
| 24   | .958    | 1.00 | 1.00 | .80 | 1.00| 48   | .979    | 1.00 | 1.00 | 1.00 | .90|
To summarize the results of the item analysis across all groups, the following observations are made:

**Visual and Motor Items** * Items linked with clutch tasks were relatively easy items for all groups;  
* Landolt rings linked with throttle tasks, and arrows linked with brake tasks were relatively harder for all groups.

**Visual Items** * Letter items were very easy test items for all groups;  
* Landolt rings were relatively difficult test items for all groups.

**Motor Items** * Clutch test items were relatively easy for all groups;  
* Brake and throttle were relatively difficult test items for all groups.

In the second phase of the item analysis, results were examined by group. The following observations were made (Refer to Table 7):

**Visual and Motor Task** * ER Group: 37 items showed perfect scores; difficulty with only 3 items;
* FTN Group: 34 items showed perfect scores; difficulty with 6 items (below .80 correct) which were all Landolt rings with throttle items, or arrows with brake items;

* NOV Group: 11 out of these 13 items (below .80 correct) were Landolt rings with throttle items or arrows with brake items;

* NM Group: 29 items showed perfect scores; 7 out of 10 items (below .80 correct) were Landolt rings with throttle items, or arrows with brake items.

Next, the following analysis of proportions on the visual task items were observed for each group (Refer to Table 8):

<table>
<thead>
<tr>
<th>Visual Items</th>
<th>ER Group: No errors were made.</th>
</tr>
</thead>
<tbody>
<tr>
<td>* FTN group: 40 items showed perfect scores; 7 of the 8 errors observed were Landolt rings or arrows;</td>
<td></td>
</tr>
<tr>
<td>* NOV Group: 36 items showed perfect scores; out of 12 error items, most were made on</td>
<td></td>
</tr>
</tbody>
</table>
Landolt rings (6) or arrows (4);

* NM Group: 43 items showed perfect scores;
out of 7 error items, 4 were Landolt rings.

Finally, the following analysis of proportions on the
motor task were observed for each group (Refer to Table 9).

**Motor Task Items**

* ER Group: 36 items showed perfect scores;
of the 12 error items, 6 were brake errors and the remainder were distributed
evenly among clutch, horn and throttle;

* FTN Group: 35 items showed perfect scores;
of the 13 error items, there was no pattern to the distribution.

* NOV Group: 18 items showed perfect scores;
the remaining items were distributed as follows: Throttle: 7 items.
  Clutch: 5 items.
  Brake: 9 items.
  Horn: 9 items.

* NM Group: 26 items showed perfect scores;
the remaining items were distributed as follows: Throttle: 7 items.
  Clutch: 2 items.
  Brake: 5 items.
  Horn: 8 items.
To recap the item analysis for groups, the following observations are made:

* ER: Highest perfect item scores, no visual errors; only significant motor errors were made on brake items.

* FTN: Few visual item errors except on Landolt rings or arrows; most motor errors on brake items.

* NOV: Most visual errors with Landolt rings or arrows; poorest overall item performance.

* NM: Problem items were Landolt rings or arrows; other error items were evenly distributed.

To summarize the section on item analysis, it was observed that Landolt ring items, arrow items, brake items, and throttle items had poorer performance; letter stimuli as test items had the best overall performance. The Experienced Riders decidedly were the best performers on all items, followed by the Formally Trained Novices, Non-Motorcyclists and Novices. Tests on item analysis and face validity have been conducted and shown to have discriminated among subjects and among groups, but not well. Furthermore, a given item was
seen to be as relatively valid as the next -- no one item performed very poorly, and all items performed above 75 per cent and most above 90 per cent.

**Test Reliability**

Anastasi (1969) says that test reliability "always means consistency" in a test. In other words, can it obtain a sufficiently consistent score from the same person at a different time? While equivalent-form or split half reliability often is used to assess test reliability over time, an odd-even reliability measurement was applied to these data for two reasons. First, there was a significant amount of task learning that would depress the reliability coefficient in a split half test, or raise retest scores or an alternative form reliability. Second, there was no reasonable way to locate the test subjects for retesting. They were widely scattered around the state. Therefore, an odd-even reliability test was applied that was not sensitive to task learning and could be applied with a single test of each subject.

An odd-even reliability test was performed on the data. The results show that the Visual Task items were correlated rather poorly ($r = .25329$), while the Motor Task items were correlated somewhat better ($r = .41631$). The Reaction Time
task items were correlated quite well ($r = .94495$). The reliability coefficients for Visual and Motor tasks were somewhat low, which is in line with their inability to discriminate very well among groups, as will be shown later in this chapter.

Because the calculation of odd-even reliability coefficients are based on half a test, they necessarily underestimate the true reliability of the test. The Spearman Brown prophecy formula (see Downie and Heath, 1974; p. 238) adjusts these coefficients upwards as if the test length were increased. On the Visual Task, the adjusted coefficient of reliability for a test doubled in length would be .40420. For the Motor Task, the adjusted coefficient of reliability for a test doubled in length would be .58788. For the Reaction Time Task, the adjusted coefficient of reliability for a test doubled in length would be .97170. Tripling the length for the first two tasks would raise their odd-even reliability further, to .60630 and .88182.

**Task Learning**

Most scientific studies attempt to control task learning to prevent diluting the results. These studies generally will allow subjects to practice the task until a skill-plateau is reached. Then, actual testing begins.
This study was constructed specifically so that all participants had to learn a new task. (See Hypotheses, Chapter I). In this way, no subject had a controls or stimulus familiarity advantage on task performance. In order to determine whether task learning was evident, an analysis was performed which compared the first ten to last ten proportions of correct responses. Table 9 reveals that the desired task learning was in evidence for each task type. Null hypotheses (a), (b), and (c) are rejected.

Table 9

Comparison of Mean Correct Responses, First Ten vs. Last Ten Items, All Groups

Visual Task

First Ten Items: $\bar{X} = .965$

Last Ten Items: $\bar{X} = .995 \quad (p = .01162)$

Motor Task

First Ten Items: $\bar{X} = .918$

Last Ten Items: $\bar{X} = .965 \quad (p = .0103)$

Reaction Time Task

First Ten Items: $\bar{X} = 1.679$

Last Ten Items: $\bar{X} = 1.081 \quad (p = .0001)$
As shown in Table 9, task learning is taking place on all three dependent variables. This suggests that the control procedure was effective. It required all subjects to learn the novel experimental apparatus. The table suggests that no single task was mastered better than another.

A follow-up analysis, by subject group, was performed on task learning. These paired t-tests revealed that all four groups performed identically. On the Visual and Motor Tasks, each group improved its scores, but not significantly. On the Reaction Time Task, each group performed significantly faster on the last ten items. Evidently, no single group had a pre-test advantage on control or stimulus familiarity. Table 10 illustrates these findings.
Table 10

Comparing First Ten Scores to Last Ten Scores by Group

<table>
<thead>
<tr>
<th>Visual Task</th>
<th>Mean*</th>
<th>Std. Dev.</th>
<th>t value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER: First Ten Items</td>
<td>.01001</td>
<td>.03166</td>
<td>1.0000</td>
<td>.3306</td>
</tr>
<tr>
<td>Last Ten Items</td>
<td>.00000</td>
<td>.00000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTN: &quot; &quot; &quot;</td>
<td>.02003</td>
<td>.04223</td>
<td>.9694</td>
<td>.3452</td>
</tr>
<tr>
<td>&quot; &quot; &quot;</td>
<td>.00589</td>
<td>.01861</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOV: &quot; &quot; &quot;</td>
<td>.05019</td>
<td>.07107</td>
<td>1.6137</td>
<td>.1240</td>
</tr>
<tr>
<td>&quot; &quot; &quot;</td>
<td>.01177</td>
<td>.02482</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM: &quot; &quot; &quot;</td>
<td>.06094</td>
<td>.12847</td>
<td>1.5000</td>
<td>.1510</td>
</tr>
<tr>
<td>&quot; &quot; &quot;</td>
<td>.00000</td>
<td>.00000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Motor Task

<table>
<thead>
<tr>
<th>ER: First Ten Items</th>
<th>Mean*</th>
<th>Std. Dev.</th>
<th>t value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last Ten Items</td>
<td>.04017</td>
<td>.07030</td>
<td>1.4908</td>
<td>.1533</td>
</tr>
<tr>
<td>FTN: &quot; &quot; &quot;</td>
<td>.03015</td>
<td>.06789</td>
<td>.2810</td>
<td>.7819</td>
</tr>
<tr>
<td>&quot; &quot; &quot;</td>
<td>.02354</td>
<td>.03039</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOV: &quot; &quot; &quot;</td>
<td>.13126</td>
<td>.11765</td>
<td>1.5966</td>
<td>.1278</td>
</tr>
<tr>
<td>&quot; &quot; &quot;</td>
<td>.06487</td>
<td>.05876</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM: &quot; &quot; &quot;</td>
<td>.13312</td>
<td>.17667</td>
<td>1.9097</td>
<td>.0722</td>
</tr>
<tr>
<td>&quot; &quot; &quot;</td>
<td>.02356</td>
<td>.04121</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 10, Continued

**Reaction Time Task**

| ER: First Ten Items | .95760 | .16171 |
| ER: Last Ten Items  | .76829 | .16895 |
| FTN: " " "          | 1.52380 | .42984 |
| FTN: " " "          | 1.02336 | .27577 |
| NOV: " " "          | 2.32280 | 1.03407 |
| NOV: " " "          | 1.33765 | .54791 |
| NM: " " "           | 1.91340 | .71747 |
| NM: " " "           | 1.19506 | .37746 |

2.5597  .0197
3.0989  .0062
2.6621  .0159
2.8020  .0118

* Means expressed as arc-sin transformations
Table 10 shows that all groups learned the experimental procedures uniformly, and no one group was significantly faster to learn the task than another. Thus, this probably did not account for observed between-group results because there are no between-group differences.

**Group Means**

Using means as measures of central tendency, it is useful to present grouped data to get the general, unspecific overview of differences before hypothesis testing. Table 11 shows group mean scores on each of the three dependent variables: Visual Task (in mean percent error), Motor Task (in mean percent error), and Reaction Time in mean seconds, all per 48 trials.
Table 11
Mean Task Scores by Subject Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Dependent Variable</th>
<th>Mean</th>
<th>Total Errors</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>Visual</td>
<td>.2000</td>
<td>2</td>
<td>.00659</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>1.8000</td>
<td>18</td>
<td>.02154</td>
</tr>
<tr>
<td></td>
<td>R/T</td>
<td>.8464</td>
<td>-</td>
<td>.16321</td>
</tr>
<tr>
<td>FTN</td>
<td>Visual</td>
<td>.7293</td>
<td>7</td>
<td>.017156</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>1.5000</td>
<td>15</td>
<td>.02824</td>
</tr>
<tr>
<td></td>
<td>R/T</td>
<td>1.2129</td>
<td>-</td>
<td>.32946</td>
</tr>
<tr>
<td>NOV</td>
<td>Visual</td>
<td>1.2000</td>
<td>12</td>
<td>.03077</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>4.3000</td>
<td>43</td>
<td>.05330</td>
</tr>
<tr>
<td></td>
<td>R/T</td>
<td>1.6027</td>
<td>-</td>
<td>.56881</td>
</tr>
<tr>
<td>NON</td>
<td>Visual</td>
<td>.7000</td>
<td>7</td>
<td>.02609</td>
</tr>
<tr>
<td></td>
<td>Motor</td>
<td>3.2000</td>
<td>32</td>
<td>.08476</td>
</tr>
<tr>
<td></td>
<td>R/T</td>
<td>1.4764</td>
<td>-</td>
<td>.51025</td>
</tr>
</tbody>
</table>

Standard deviations appear as arc-sin transformations for Visual and Motor Task in order to improve normalcy.
Assumptions of the Analysis of Variance

Downie and Heath (1974) discuss the following assumptions of analysis of variance, a test often used in hypothesis testing:

1) The individuals in the various subgroups should be selected on the basis of random sampling from normally distributed populations.

2) The variance of the subgroup should be homogeneous.

3) The samples that make up the groups should be independent (p. 41).

These authors continue as follows; "It has been shown that the analysis of variance technique is robust in respect to these assumptions. This means that when the analysis of variance is used, the results will be accurate even if the homogeneity assumption is violated. However, sample sizes should be the same or very similar in number. Likewise, the assumption of normality of distribution may be violated, providing the departure from normal is not too large" (Downie and Heath, 1974, p. 274).

An examination of the assumptions for the analysis of variance follows:

1) As volunteers, these individuals in the groups ER, FTN, NOV, and NM were not randomly selected but came from
large and assumedly normally distributed parent populations.

2) The assumption for equal variances was violated in this study as suggested in Table 12. Table 12 uses only the dependent variable reaction time since it is the only available continuous dependent variable. Even with arcsin transformations to make the discreet variables (Visual and Motor Task) continuous and more normally distributed, there were still too many perfect scores on these tasks, rendering a large amount of "zero variance" subjects, and rendering their variance comparisons unintelligible. A "ceiling effect" was apparently taking place on these two tasks.

Table 12

<table>
<thead>
<tr>
<th>Group</th>
<th>Variance on R/T</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>.02663</td>
</tr>
<tr>
<td>FTN</td>
<td>.10854</td>
</tr>
<tr>
<td>NOV</td>
<td>.32355</td>
</tr>
<tr>
<td>NM</td>
<td>.26035</td>
</tr>
</tbody>
</table>

As Hartley's test (1950) for homogeneity on variance shows, the F-max statistic exceeds the critical value of 6.31.

\[
F_{\text{max}} = \frac{\text{largest}}{\text{smallest}} = \frac{.32355}{.02663} = 12.150
\]

\[
df=(4,9) \quad CV=6.31
\]
Therefore, the hypothesis is rejected that there is a significant difference in group variances.

3) Samples comprising the groups are mutually exclusive and subjects do not serve under two treatments— they are drawn independently.

Boneau (1960) and Norton (1952) have shown that when violations of normalcy and homogeneity of variance occur, too many or too few significant differences appear, resulting in type I and type II errors. All authors urge caution in subsequent statistical procedures, yet many note that it is impossible to anticipate some of these violations, especially heterogeneity, unless there is a large body of literature from which to anticipate group variances (Wike, 1971). Unfortunately, motorcyclists' human factors literature lacks a body of experimental research from which to anticipate heterogeneity of variance.

Authors are divided on how to proceed when marked heterogeneity has been demonstrated. On the other hand, certain statistical treatments have been suggested which do not rely on the assumptions of normalcy, independent observations and homogeneity of variance. These statistical procedures are known as distribution-free, or non-parametric tests.
On the other hand, other authors rely on t and F tests because they view these parametric tests as inherently strong, even in view of violations of assumptions (see Lindquist; Box 1953). One of these authors, Box, has termed parametric tests "robust" because they "are only inconsequently affected by a violation of the underlying assumptions." (Bonferroni, 1960 p. 61) Rather than abandoning parametric analyses, if heterogeneity is observed, they suggest using more stringent criteria: higher significance levels, for example (Lindquist, 1953, p. 86).

In this research on the effects of training on novice motorcyclists, it was decided to employ both types of tests, for multiple reasons. First, it could not be anticipated that violations of homogeneity of variance would occur because of a lack of baseline data from which to make these inferences. Second, in breaking new ground as this study does, and in light of the problems encountered with the test instrument itself, it probably is best to analyze the data so the effects of the treatment are not overlooked. It was decided to first perform a non-parametric test to determine if differences between the two groups, trained and untrained novices, are evident. If significant differences were found, then "robust" parametric tests would be undertaken, and that includes the analysis of variance. Finally, it can be easily
shown that the elimination of the experienced riders (ER) and the non-motorcyclists (NM) leaves the two groups of importance (FTN and NOV) with relatively homogeneous variances ($F_{max} = 2.9809$). This adds utility and validity to the use of parametric comparisons of these two groups.

One commonly-used two-group non-parametric test is the Mann-Whitney test. Applying this test to the FTN and NOV groups, the following observations were made. On the Visual Task, no significant differences were observed ($p = .218$; $m=10$, $n=10$). On the Reaction Time Task, no differences were found ($p = .062$; $m=10$, $n=10$). However, there were significant differences between groups on the Motor Task ($p = .001$; $m=10$, $n=10$). This finding suggests that trained riders react more appropriately on a motorcycle handlebar simulator than untrained riders, an important finding considering Hurt's conclusions that motorcyclists involved in accidents tend to make confused, inappropriate responses. These findings also suggest a more complete analysis, including parametric tests, despite the violations of certain underlying assumptions.

Prior to undertaking parametric tests, it would be worthwhile to anticipate the other possible sources of the observed variance. One way to do so is by locating "outliers", and another way is to perform a gender analysis suggested by
Table 5, which shows that more females were found in two experimental groups than would be expected in real life. In investigating "outliers", a spreadsheet of subject scores was created to find any subjects who might account for much or most of the variance. See Table 13.
Table 13

Subject Error Scores by Subject Group

<table>
<thead>
<tr>
<th>ER</th>
<th>Visual</th>
<th>Motor</th>
<th>R/T (avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1.095</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0.657</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0.685</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0.917</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0.715</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>4</td>
<td>0.762</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>3</td>
<td>1.092</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>2</td>
<td>0.884</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0.746</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>2</td>
<td>0.937</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NOV</th>
<th>Visual</th>
<th>Motor</th>
<th>R/T (avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>4</td>
<td>8</td>
<td>2.525</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>2</td>
<td>1.190</td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>3</td>
<td>1.460</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>2</td>
<td>1.114</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>9</td>
<td>0.810</td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>3</td>
<td>1.602</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>4</td>
<td>1.212</td>
</tr>
<tr>
<td>28</td>
<td>3</td>
<td>6</td>
<td>1.888</td>
</tr>
<tr>
<td>29</td>
<td>2</td>
<td>4</td>
<td>1.763</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>2</td>
<td>2.444</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FTN</th>
<th>Visual</th>
<th>Motor</th>
<th>R/T (avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>2</td>
<td>1</td>
<td>1.398</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>1</td>
<td>1.319</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1.505</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0</td>
<td>0.813</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>1</td>
<td>0.960</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>2</td>
<td>1.202</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>0</td>
<td>1.034</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>2</td>
<td>0.963</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>3</td>
<td>1.097</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>4</td>
<td>1.907</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NM</th>
<th>Visual</th>
<th>Motor</th>
<th>R/T (avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>1</td>
<td>13</td>
<td>2.083</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>2</td>
<td>2.342</td>
</tr>
<tr>
<td>33</td>
<td>3</td>
<td>2</td>
<td>0.911</td>
</tr>
<tr>
<td>34</td>
<td>3</td>
<td>3</td>
<td>1.114</td>
</tr>
<tr>
<td>35</td>
<td>0</td>
<td>0</td>
<td>1.088</td>
</tr>
<tr>
<td>36</td>
<td>0</td>
<td>6</td>
<td>0.955</td>
</tr>
<tr>
<td>37</td>
<td>0</td>
<td>0</td>
<td>1.680</td>
</tr>
<tr>
<td>38</td>
<td>0</td>
<td>0</td>
<td>1.780</td>
</tr>
<tr>
<td>39</td>
<td>0</td>
<td>5</td>
<td>1.728</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
<td>1</td>
<td>1.082</td>
</tr>
</tbody>
</table>
Clearly, the NM group is widely varied as indicated by its broader variance, and also the outliers in subjects 31, 36, and 39. The outliers in the NOV group were subjects 21, 25, and 28. Further tests are called for to locate the actual source of variation.

**Gender Differences**

Next, in an attempt to pinpoint other sources of variance, a gender analysis was performed which compared male to female performance, recalling that two groups, the NOV and NM groups, were weighted with higher proportions of female subjects than might be representative of "real life".

To rule out the hypothesis that gender differences accounted for any observed results, paired t-tests were applied to the raw data as illustrated in Table 14. This compares male a female scores with all tasks combined. There were significant male-female differences on Reaction Time ($p = .0009$).
Table 14
Comparing Male to Female Task Scores, by Task

<table>
<thead>
<tr>
<th>Task</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>t Value</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males: 18</td>
<td>18</td>
<td>.00936*</td>
<td>.01783</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females: 22</td>
<td>22</td>
<td>.01985</td>
<td>.02566</td>
<td>-1.3546</td>
<td>.1835</td>
</tr>
<tr>
<td>Motor Task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males: 18</td>
<td>18</td>
<td>.04407</td>
<td>.04536</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females: 22</td>
<td>22</td>
<td>.06653</td>
<td>.06317</td>
<td>-1.2644</td>
<td>.2138</td>
</tr>
<tr>
<td>Reaction Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males: 18</td>
<td>18</td>
<td>1.0045</td>
<td>.28691</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females: 22</td>
<td>22</td>
<td>1.4804</td>
<td>.52082</td>
<td>-3.4649</td>
<td>.0009</td>
</tr>
</tbody>
</table>

* Means expressed as arc-sin transformations.
Based on this table, a further comparison by gender of subject groups was performed. This time, almost no differences were found. (The only significant difference was observed for the Visual Task in the Experienced Rider group (ER), but it must be noted that this group had only a single subject and the results are not meaningful.) See Table 15.

Table 15

Comparing Male to Female Task Scores by Group

(Probability Levels)

<table>
<thead>
<tr>
<th>Group</th>
<th>Visual Task:</th>
<th>Motor Task:</th>
<th>Reaction Time:</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER:</td>
<td>p = .0353</td>
<td>p = .8655</td>
<td>p = .3096</td>
</tr>
<tr>
<td>FTN:</td>
<td>p = .7245</td>
<td>p = .0780</td>
<td>p = .9595</td>
</tr>
<tr>
<td>NOV:</td>
<td>p = .2167</td>
<td>p = .4856</td>
<td>p = .0693</td>
</tr>
<tr>
<td>NM:</td>
<td>p = .3412</td>
<td>p = .9192</td>
<td>p = .6549</td>
</tr>
</tbody>
</table>

On the whole, groups’ Reaction Time scores are affected by gender. Males and females performed with the same accuracy, but not the same speed. By groups, the differences are not detectable, probably due to the small sample sizes.
Multivariate Results

Because the non-parametric analysis showed that there may be some effects of the experimental treatment, it was decided to proceed with parametric tests. The first test was performed on an analysis of all three dependent variables, followed by an analysis of each variable separately.

A multivariate analysis of variance (MANOVA) was calculated using three dependent variables together. The multivariate null hypothesis (H) states that there is no statistically significant difference in Visual Task, Motor Task and Reaction Time considered together in a simulated motorcycle performance tests among groups of trained, untrained, experienced and non-motorcyclists. The results of this test are found in Table 16, and reveal that, overall, there are significant differences among the four groups taken together. The F value of 2.55 (df = 9, 98) exceeds the critical value of 1.96 and suggests that the probability of these results occurring by chance alone are small (p = .0114). Thus, the null hypothesis (H) is rejected, and the alternative hypothesis, that there are significant differences present, becomes tenable. There are significant differences among groups on the Visual Task, Motor Task, and Reaction Time Task when considered all together (See Table 16).
Table 16

Multivariate Summary Table

Hotelling-Lawley Trace = .70133

df = (9,96); p = .0114

Critical Value = 1.96

The Analyses of Variance

Because the multivariate analysis of variance showed significance, a series of univariate analyses of variance were undertaken. When using a single dependent variable and applying it to more than two groups, an analysis of variance provides answers to the hypothetical inference that there is no difference among group means.

These symbolic expressions are the statistical equivalents of the following hypotheses:

e) There are no statistically significant differences among ER, FTN, NOV, and NM groups on the dependent variable, Visual Task error scores, as measured by the ANOVA;

f) There are no statistically significant differences among ER, FTN, NOV, and NM groups on the dependent variable, Motor Task error scores, as measured by the ANOVA;

g) There are no statistically significant differences among ER, FTN, NOV, and NM groups on the dependent variable,
Reaction Time, as measured by the ANOVA.

**Visual Task ANOVA Results**

The SAS computer assisted analysis packages provide a general linear model procedure for the analysis of variance (ANOVA). Table 15 displays the ANOVA results on the test of the Visual Task over all groups. As can be noted, the F value of 1.79 does not exceed the critical value of 2.86 (df=3,36) and the probability of these results occurring by chance alone are relatively high (p=.1663): thus, the null hypothesis (e) is not rejected. There are no statistically significant differences among groups on the Visual Task. See Table 17.
Table 17

Univariate Summary Table, Visual Task Errors

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3</td>
<td>0.00264037</td>
<td>0.00088012</td>
<td>1.79</td>
<td>0.1663</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>0.01768933</td>
<td>0.00049137</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected total</td>
<td>39</td>
<td>0.02032970</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Critical Value F (3,36)= 2.86
Motor Task ANOVA Results

As can be seen in Table 18, ANOVA results are examined for the Motor Task error scores among all four groups. The derived F value of 2.62 does not exceed the critical value of 2.86 (df = 3,36) and the probability of these results occurring by chance alone are relatively low (p = .0654). Thus, the null hypothesis (f) is not rejected. There are no statistically significant differences among groups on the Motor Task, but a more reliable test might have detected differences. In that case, the application of Lindquist's suggestion of a more stringent criteria (e.g. alpha = .01) might have also overcome problems of heterogeneity. See Table 18.
<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3</td>
<td>.02219942</td>
<td>.00739981</td>
<td>2.62</td>
<td>.0654</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>.10157862</td>
<td>.00282163</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>39</td>
<td>.12377804</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Critical Value $F(3, 36) = 2.86$
Reaction Time ANOVA Results

In Table 19, ANOVA results can be seen for the significance test on the Reaction Time task for all groups. The derived F value of 6.21 exceeds the critical value of 2.86 (df= 3,36) and the probability of these results occurring by chance are relatively small (p=.0016). Thus, the null hypothesis (H0) is rejected, and its alternative hypothesis is accepted. There are significant differences among the four groups on the Reaction Time Task. The highly significant finding on Reaction Time suggests that it is stable, and not affected by heterogeneity. See Table 19.
Table 19: Univariate Summary Table, Reaction Time

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F Value</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>3</td>
<td>3.35075168</td>
<td>1.11691123</td>
<td>6.21</td>
<td>0.0016</td>
</tr>
<tr>
<td>Error</td>
<td>36</td>
<td>6.47178264</td>
<td>0.97977174</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>39</td>
<td>9.82253432</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Critical Value $F(3, 36) = 2.86$
**Between Groups Comparisons**

Tables 17 and 18 show that there were no significant differences among groups on the Visual Task or Motor Task, but significant differences did exist among groups on the Reaction Time task, as seen in Table 19. Therefore, a test of paired comparisons was applied to the Reaction Time scores to determine exactly where the differences appeared. Table 20 shows the results of a Tukey's Standardized Range (HSD) test, applied at alpha = .05. Table 20 suggests that there are no significant differences in reaction times between the ER and FTN groups. There are significant differences observed between the ER and NOV groups and the ER and NM groups, but since Table 14 shows that males and females have significantly different reaction times, the observed differences in Table 20 could be due to male-female differences alone. Table 20 follows.
Table 20

Pairwise Comparisons for Reaction Time Task

<table>
<thead>
<tr>
<th>Grouping</th>
<th>Mean (sec.)</th>
<th>N</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.4764</td>
<td>10</td>
<td>NM</td>
</tr>
<tr>
<td>A</td>
<td>1.6027</td>
<td>10</td>
<td>NOV</td>
</tr>
<tr>
<td>A</td>
<td>B A</td>
<td>10</td>
<td>FTN</td>
</tr>
<tr>
<td>B</td>
<td>.84665</td>
<td>10</td>
<td>ER</td>
</tr>
</tbody>
</table>

alpha=.05 (all Means with the same letter under "Grouping" are not significantly different).

Correct vs. Incorrect Reaction Times

A follow-up analysis was done to determine whether incorrect Visual or Motor Task response reaction times were substantially longer than the correct response times. Since the ER or FTN groups were more familiar with motorcycle controls, they might be less confused if they knowingly made an incorrect response. If this phenomenon occurred, then groups with little or no experience might have faster reaction times on the whole and thus influence findings on the parametric tests. However, the analysis in Table 21 shows that this did not occur. The paired t-tests show that no group had significantly faster reaction times on incorrect Visual or Motor Task reaction times; there was no detectable
residual effect of making a mistake on the reaction time of the next test item.

Table 21

Mean Response Time Difference Between Correct and Incorrect Responses, By Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Difference (in 1/100 sec.)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER</td>
<td>.46</td>
<td>.6544</td>
</tr>
<tr>
<td>FTN</td>
<td>-1.53</td>
<td>.1613</td>
</tr>
<tr>
<td>NOV</td>
<td>-1.50</td>
<td>.1666</td>
</tr>
<tr>
<td>NM</td>
<td>-1.62</td>
<td>.1400</td>
</tr>
</tbody>
</table>

Discussion of Research Results

This chapter has presented the results of research on whether trained and untrained motorcyclists as well as experienced and non-motorcyclists differ on Visual, Motor and Reaction Time tasks. The simulator used in this research is at least face-valid because it is an actual handlebar and controls from a full-size motorcycle.

This research had relatively low internal validity which was reflected in low reliability coefficients on the Visual and Motor Task. Doubling or tripling the test's length would, according to the Spearman-Brown Prophecy formula, bring the reliability up to more acceptable levels, raising the test's
internal validity and perhaps showing significance on some results with marginal probabilities, such as the Motor Task (p = .0654). Particularly low item scores were not observed for any Visual Task items when considered together with the Motor Task items. However, on the Visual-only Task Landolt rings were more difficult items, as were arrow items. On the motor-only task, throttle items and brake items were more difficult items. Most items were at a .90 proportion correct, signifying a ceiling effect: a test that was "too easy". For reliability, a low odd-even correlation coefficient on Visual Task, and on Motor Task, was observed, indicating a test that did not differentiate well. A high coefficient was observed for Reaction Time. A Hartley's F-max test showed that the assumption of homogeneity of variance was violated, probably due to outliers who had perfect scores, and some who had very poor scores.

On the whole, subjects demonstrated the desired task learning, judged by significantly lower scores on the last ten items of the 48 trials. Task learning was in evidence on each task overall, but a task learning analysis by group showed that only Reaction Times were significantly faster on the last ten test items. This was probably a function of small samples, poor test reliability or the ceiling effect on the Visual and Motor Task: these tasks didn't show increased
proficiency by the group analysis presented in Table 10.

A non-parametric test suggested that there were Motor Task treatment effects on the FTN group which demonstrated significantly fewer errors. This suggested proceeding to the use of parametric tests, despite the violation of some underlying assumptions about the use of parametric tests, relying on increased test stringency instead. A multivariate analysis of variance showed significant differences when considering all three dependent variables simultaneously. Following this, three univariate analyses of variance were performed. The analysis of variance showed that neither the Visual Task nor the Motor Task discriminated among groups significantly but that Reaction Times did discriminate significantly among groups. The training effect seen in the Mann-Whitney test for FTN subjects must be replicated to confirm its true nature as a stable or transitive phenomenon. A test on pairwise comparisons for the only significant variable, Reaction Time, showed that ER and FTN groups were not significantly different, but ER and NOV, and ER and NM groups were significantly different on the Reaction Time measure.

In an effort to support and justify use of parametric tests, Lindquist has suggested more stringent significance levels, e.g., .01. In this research, the multivariate analysis of variance and the univariate analysis of variance on the Reaction Time task met this criterion, and on the basis of
Lindquist's suggestion, may be considered truly significant.

A follow-up analysis investigated whether response times were longer or shorter when a subject made an error on the Visual or Motor task. However, as Table 21 showed, there was no detectable significant difference between correct and incorrect responses for any subject group. There was no evidence of a "residual" effect of making a mistake.

However, as mentioned, quasi-experimental research designs often are subject to alternate explanations (rival hypotheses) to account for observed results. Four such alternate hypotheses are dealt with here.

**Pretest Differences in Control Familiarity.** Pretest differences in subject familiarity with motorcycle controls was partially controlled by requiring all subjects to learn a new untypical experimental task that did not just depend only on the amount of motorcycle operator experience. This control was discussed under Task Learning, where Table 9 revealed that all subjects had significantly better Visual and Motor Task scores and faster Reaction Times on the last ten items compared to the first ten. In addition, Table 10 showed that no single group learned any task faster than another group. The rival hypothesis that pre-test control led to faster or more accurate scores (by task or by group) can be tentatively
Representativeness of Groups. The ER and FTN groups had approximately the same proportion of males and females as is represented in real life (Motorcycle Safety Foundation, 1984):

ER: 90 percent Male, 10 percent Female;
FTN: 50 percent Male, 50 percent Female.

However, the remaining two groups were not representative of real life; both the NOV and NM groups were overly represented by females. However, two analyses (illustrated in Tables 14 and 15) showed that there were no between-group gender differences even though there were overall gender differences on Reaction Times. This rival hypothesis can't be dismissed.

Small Samples. Only a single, small group of 10 FTN test subjects were available for testing between January 1, 1985 and May 1, 1985. Small samples tend to place emphasis on outliers (Table 12) and this has shown up in the relatively large group variances observed. Larger groups tend to be more normally distributed according to the Central Limit Theorem, reducing variation and increasing normalcy. The rival hypothesis that small sample sizes accounted for observed dependent variance scores cannot be dismissed, and probably represents a weakness of the study.
The Ceiling Effect. It will be recalled that the univariate ANOVA did not find differences among groups on either the Visual or Motor Task. In addition, the item analysis (pp. 90-100) revealed that many test items had perfect scores all the way through. The important point here is not that the Visual and Motor Tasks were not significant, but that they did not discriminate well enough to confirm or disconfirm the implicit hypothesis about training. The competing explanation here is that a different test, i.e., a more difficult Visual and Motor Task, would have reduced this skewness and shown real differences. Such a test would discriminate better, and probably have improved test reliability as a consequence. This rival explanation cannot be dismissed, either.

Despite the violations of certain assumptions of parametric tests (t and F), two reasons point to the validity of these tests. First, non-parametric tests (Mann-Whitney) revealed a training advantage for the FTN group. This test significant results, and suggested further data examination by way of parametric tests, the application of more stringent significance levels. On this second test, the multivariate analysis of variance and the reaction time analysis were still significant, but not in regard to training. Larger samples, as mentioned, would be the single most likely way to crosscheck these results by way of replication.
In summary, it can be seen that neither the Visual Task nor the Motor Task were good discriminators, and the results shown are not meaningful because the rival hypotheses vitiate any clear interpretation. More difficult Visual and Motor Tasks and a longer test would reduce the ceiling effect, raise test reliability, and may yet discriminate. However, the Reaction Time Task performed much better, showing a significant univariate analysis of variance (Table 17). We see that the Reaction Time Task did discriminate among groups, but not on the training available. These results are more interpretable and hence, more useful, because of the high test reliability shown. Though results do not say however that training has the expected significant effect, although future research may yet discover this.

Chapter V will present recommendations based on the research results presented in Chapter IV. Some guidelines for future research are also presented.
Chapter V

CONCLUSIONS AND RECOMMENDATIONS

This chapter ties together the rationale for the project with its findings, and draws specific inferences within the limitations of the research methods. It does this by first reviewing the rationale for the research, including the problem statement; then it proceeds to review briefly the relevant literature that led to the project; then this chapter reviews the study's methods and research procedures; finally, this chapter reiterates the research results generated within its specific parameters. The purpose of this chapter, then, is to bind together the previous four chapters with legitimate generalizations on its findings. It will make recommendations and inferences that have value in the real world of highway safety research, but which are tempered by the restrictions of statistical inference, and research constraints.

Review of the Problem

The problem of motorcycle accidents is well documented in literature on highway safety. Chief among the causes of these accidents are operator inexperience and operator error. These show up when motorcycle operators make inappropriate or slow responses to a potentially dangerous situation.
Driver educators, legislators, motorcycle enthusiasts and the motorcycle industry propose motorcycle training, but there is no established empirical link between the training and the improved response accuracy and quicker reaction times. Previous research shows novice drivers scan ineffectively and that motorcyclists have trouble "time-sharing" tasks with competing demands.

Review of the Rationale of The Study

The rationale behind this investigation is that training should improve the static visual acuity, motor skills and reaction times of motorcycle operators. Yet the literature review shows that there has been very little basic research on the human factors aspect of motorcycle operation and so experimental parameters are unclear since there is little to build on.

Even though there have been no useful human factors studies on the perceptual processes of motorcyclists, this study attempts to use the findings of other research on auto drivers, and add a simultaneous motor task with the visual task. In this way, this research attempts to eliminate one of the methodological problems shown in past research by adding a secondary task that should demonstrate "time-sharing."
Finally, it is suggested that a unifying paradigm can be useful when applying to real life any findings from human factors research in automobile and motorcycle operation because paradigms offer a vehicle for interpreting and organizing the findings. Information processing is a good candidate for that research paradigm.

A research project was designed which compared the performance of trained to untrained motorcyclists on three dependent variables. Experienced Riders and Non-Motorcyclists were also added to extend the findings and offer a control group. All four groups were tested on a Visual Task, a Motor Task and Reaction Time. A motorcycle simulator was used to test these reactions in a laboratory situation which controlled ambient conditions and provided a stress-free, non-traffic environment.

Review of Methods and Procedures

A quasi-experimental design was developed which compared Experienced Riders, Formally Trained Motorcyclists, Novice Motorcyclists and Non-Motorcyclists on the basis of three dependent variables, namely visual, motor, and reaction time measures. Volunteers were drawn from available populations including 10 formally trained students, which had elected to
participate in a motorcycle training course at the Cuyahoga Community College during the last three weeks in April of 1985. Subjects viewed and identified four types of visual stimuli and responded to them on a motorcycle handlebar simulator. Responses were recorded for each of 48 trials, and a total of 144 responses were recorded for each subject. A pilot study, using 16 pilot subjects (four in each category), showed that the implicit hypothesis was supported and that minor procedural aspects could be changed to facilitate a full research project.

Review of Research Findings

An analysis of the reliability and validity of the test instrument turned up a problem with the test instrument's reliability, probably due to the low difficulty of the Visual and Motor Tasks. This ceiling effect was demonstrated by item analysis. In addition, a test for homogeneity of variance on Reaction Time scores showed that the four groups had significantly different variances. This could be due to the rather small sample size and possibly by gender differences. Because of these problems in interpretation, it is not clear whether the Visual and Motor tasks showed "no differences" or simply didn't discriminate.
An item analysis also revealed that Landolt rings and arrows were relatively difficult test items compared to the more common alphabetic and numeric characters. The required task learning was exhibited, to a statistically significant degree overall and by group, as shown by significantly quicker reaction time scores on the last ten items.

The Visual Task was designed to test static acuity during simulated motorcycle operation. The Motor Task was designed to measure subjects' motor skills, e.g., how well were stimuli linked to the appropriate motor response. Reaction times were taken on each task. There were no statistically significant differences among groups on the Visual Task, and there were no statistically significant differences among groups on the Motor Task. However, the ceiling effect, or small, highly variable samples were the likely cause of these results. However, there were statistically significant differences among groups on the Reaction Time Task. In addition, there were overall statistically significant differences among groups when considering all three tasks at the same time.

A pairwise comparison test, Tukey's Standardized Range Test, was accomplished on the only significant dependent variable, Reaction Time; the ER and FTN groups weren't significantly different, but the ER and NOV, and the ER and NM
groups were significantly different. This signifies that the ER group is faster than novices and non-riders, and that the ER group is indistinguishable from the FTN group on Reaction Time, but nothing can be concluded about any advantages of training, until further research can overcome this study's shortcomings by replication.

Conclusions from the Study's Results

Based on the research results provided, these conclusions were reached:

1) Fundamental human factors research on motorcycle operators has now been accomplished under controlled laboratory conditions. In addition, the motorcycle simulator used in this study did discriminate on the Reaction Time in the direction of the hypothesis, but not to a magnitude to make inference in training. It is clear that a simulator may be a useful way to test motorcycle operators.

2) Non-parametric tests revealed a Motor Task advantage for trained riders in the FTN group; this finding prompted the use of parametric tests, even though fundamental assumptions were violated. Relying on guidance of authors who encourage the use of "robust" t and F tests in these cases, and requiring more stringent significant levels (e.g., .01),
there were significant findings detected but none that suggested an effect for formal training. The training advantage observed in the FTN group in the Mann-Whitney test at the marginal significance level on the parametric ANOVA must be confirmed by replication.

3) It is apparent that the Visual Task and the Motor Task were too easy; very few errors were made by any group in these tasks.

4) On the other hand, the Reaction Time Task was a better discriminator, showing significant differences for experienced riders compared to novices and non-riders.

5) It appears that there is still no substitute for experience, at least on the Reaction Time Task, since the results show significant differences between ER and NM, an ER and NON groups. If all results were taken only at face value, they would suggest Experienced Riders can do all the tasks better. However, the Visual and Motor Tasks didn't discriminate well enough to make this conclusion because of the ceiling effect.

6) The task learning analyses showed that the letters, numbers and abstract symbols are useful experimentally because all subject groups did learn to recognize and act appropriately upon the stimuli as the test progressed. All subjects had
to learn something new, so that no group had a pre-test advantage on control or stimulus familiarity.

7) Males do not perform these experimental tasks more accurately than females, but males do perform significantly faster. Because these findings are based on a male sample of 18 and a female sample of 22, they are probably not transitive, and would likely carry through to the group analysis if the groups themselves were larger. As it stands, male-female differences represent a confounding.

8) Incorrect responses didn't significantly affect any group's subsequent Reaction Times. There is no obvious or residual disadvantage of making a mistake, as far as Reaction Time is concerned.

9) Problems in data interpretation sprung from four alternate hypotheses discussed in Chapter IV. The task learning question has been dismissed, because task learning was a desired feature of the experiment. Problems with gender differences, the relatively small sample sizes, and ceiling effect are confounding influences and can only be dismissed through subsequent study.

10) A reliable test apparatus has been developed that will facilitate replication and further study in motorcyclist human factors.
General Recommendations from the Study's Results

Based on the research results presented in Chapter IV, and the conclusions presented above, these general recommendations may now be reached:

1) A valuable, within-subjects, repeated measures research project is recommended to overcome the confoundings described here. Large representative samples (30 or more) could be drawn from each of the same four groups; ER, FTN, NON, and NM. Pretest measures on one, two, or all three more difficult dependent measures (Visual Task, Motor Task, and Reaction Time) would be taken and compared to post-test measures. The FTN group would, of course, proceed with a standardized training course, such as the MRC described. Pretest measures should show no significant group differences, while on the post-test measures, the FTN group should improve significantly. This design minimizes the representativeness confounding, the ceiling effect and probably the low-reliability coefficients. Larger groups minimize the problem with small samples, and the within subjects, repeated measures aspect reduces the problem with between-group, pre-test individual differences. This hypothetical study could be made even more elegant by testing all subjects on each task separately, then two, then three simultaneously. All but trained
novices should show fairly "flat" curves, suggesting experience via training facilitates performance of simultaneous tasks. Such a study could better address whether motorcyclist training is a cost-effective alternative to experience.

2) Because of the observed "ceiling effect," it is recommended that future motorcycle simulations contain the full compliment of motorcycle controls. Adding motorcycle foot controls to this particular simulator would have increased the number of error choices for each operator and made time-sharing more evident. This should, at least intuitively, differentiate the groups on the Motor Tasks and Reaction Times. It is recommended that a more complete simulator is used in future research to broaden the range of observed differences, and reduce the ceiling effect.

3) Few errors were observed on the Visual Task. The Visual Task could be made more difficult by adding more letters or numbers, or by placing the Visual Tasks further into the subject's visual field. This would also broaden the range of observed differences in future research, also reducing the ceiling effect. Furthermore, rather than a static display presented by a slide projector, future researchers might consider dynamic displays in such situations in order to test cue searching and not just cue identifying. It will be recalled that Rockwell says novice drivers exhibit "frantic cue
searching" during driving and it may be that training allows the learning of cue-searching protocols. In other words, it may be that simple cue identification is not relevant as measured by the static visual acuity test in this research. A dynamic display, with moving cue targets could show whether training improves identification and also scanning techniques. A dynamic visual field would add realism and many more cues from which to search.

4) The use of non-parametric tests and also the use of more stringent significant levels in parametric tests is recommended if future research determines that widely heterogeneous groups are found upon replication. By using both types of tests, no important findings are overlooked. If heterogeneous group variances are a stable phenomenon, non-parametric tests alone are strongly recommended to underscore parametric test findings.

5) Since gender differences on Reaction Time seem to be apparent and represent a stable overall phenomenon (Table 14), further research might consider using only males or only females to obviate possible confounding.

6) It has been concluded that Formally Trained Motocyclists react faster than all other groups except Experienced Riders, suggesting that experience is still the best "training" of all. Yet gaining experience requires exposure to
potentially dangerous traffic, and in addition, gaining experience takes time. It is recommended that further research be conducted to determine at exactly what point the trade-off between experience and training occurs. This could benefit potential novices who might not be interested in training: if the risks could be demonstrated ahead of time, they might be motivated to try a training course instead.

7) It has been concluded that the finding of Novice Motorcyclists doing worse than expected on all three tasks runs against expectation. It is recommended that research be conducted to clear up the source of this apparent contradiction. Are novices "motorcycle specific," that is, do they have difficulty transferring skills once they have learned the task on a specific motorcycle? Are the results of this research confounded by the differences between groups? Are the observed conflicting differences due somehow to the ceiling effect, and would they disappear if there were no ceiling effect? Additional research is necessary to answer these questions.

8) Additional Reaction Time research with a larger, more representative sample, could determine whether the observed differences are still in the direction of the hypotheses, and show the significance of training.
9) Future researchers should avail themselves of the reliable test apparatus described herein, including the computer program and electronic interface. The cost to build this equipment is relatively low, which should foster replication.

10) Finally, it is recommended that more post-experimental debriefing be undertaken with all subjects in future research. Perhaps by orally discussing and eliciting difficulties or confusion, insight may be gained on how to improve the study's methods, or how to change procedures to highlight differences between groups relative to formal training.

Developing a Paradigm

While these results are tempered by certain confoundings, it appears that the limited findings can be interpreted in the framework of an information processing paradigm. This gives them added meaning by attaching them to observed human factors phenomena. It has been hypothesized that when a novice untrained motorcyclist is faced with a quick decision, he often makes no response or an inappropriate response to the problem. The results presented here imply that experienced riders are better able to perform Reaction Time tasks than
are novices or non-motorcyclists. The ER group can evidently "time-share" better, but whether this comes from something temporal (e.g., passage of time) or whether it comes from pre-test differences and/or sample size remains an open question. It is a tantalizing issue on exactly how the passage of time seems to differentiate experienced motorcyclists from novices, besides the obvious temporal agent.

Problems with conducting a true experiment will likely remain because of the fact that subjects self-select themselves, and true random assignment to groups is impossible, with resulting difficulties in interpretation. In short, it is premature to postulate much on information processing at this point. Further exploratory research is required to give additional insights.

Interpreting the Research Conclusions and Recommendations

Some applied human factors research in motorcycle accident avoidance has now been accomplished. This research is neither difficult, time consuming, nor expensive to conduct and should be continued as long as there is still a motorcycle accident and fatality problem. There is much to be learned about how motorcyclists perceive and react, and a basic groundwork has been laid. A laboratory controlled motorcycle simulator is a reasonable and reliable test
instrument to investigate these questions. Beyond establishing
the utility of a testing apparatus that can be used reliably
for motorcyclists, a limited number of conclusions can be
made. Chief among them is that a taxing test must be used
to improve the ability to discriminate among groups on visual
acuity and motor skills. Reaction time measures did discrim-
inate, but firm conclusions about training cannot yet be
drawn because there were no statistically significant differ-
ences between the FTN, NM, and NON groups, and because of the
male-female confounding.

Despite failing to detect differences due to training,
which may still be in evidence by other tests, this research
has progressed in measuring motorcyclist human factors phen-
onema where no previous research is available. Scientific
parameters have been established and useful suggestions
ions made for future research. For example, a motorcycle
simulator with a full complement of controls seems a useful
way to discriminate motorcyclists' reaction times. Further
research is necessary to determine whether a more difficult
Visual and Motor Task can discriminate and determine an
effect for training. Additional study will also broaden
our knowledge of information processing as an explanatory
paradigm. Unless we are insensitive to the issues of saving
lives in highway safety, this research should continue.
APPENDIX A
Computer Program for Apparatus Timers

1 REM PROGRAM DEVELOPED BY ROBERT J. VORAK
5 PB=37136: PA=3718
8 DIM D(50), E(50)
9 P=0: L=0: LP=0
10 POKE PA,128: POKE PB,255
20 PRINT "HEART"
25 INPUT "NUMBER OF SLIDES";TV
26 PRINT "HEART";
30 INPUT "SUBJECT";Z$
31 PRINT "HEART"
32 PRINT: PRINT: INPUT "HIT RETURN TO BEGIN";YCS
40 GOSUB 100
45 PRINT "HEART" "QQQQQQQ TEST IN PROGRESS"
50 TI$="000000"
60 A$=" ": GOTO 200
100 POKE PB,0
110 FOR I=1 TO 100: NEXT I
120 POKE PB, 255
130 RETURN
200 Y=PEEK (PB)
205 GET A$: IF A$<>" " THEN D(L)=500: L=0: LP=LP+1: GOTO 300
210 IF Y=P OR Y=255 THEN GOTO 200
220 D(L)=Q: E(L)=TI
230 L=L+1: P=Y
235 PRINT: PRINT "CIRCLE"L, TI
240 GOTO 400
300 PRINT "HEART"
310 PRINT "TEST # SUBJECT": PRINT LP, Z$
320 PRINT: PRINT
330 PRINT "DEVICE TIME PUSHED"
335 PRINT
340 FOR I=0 TO 1
345 H=(INT((E(I)/60)*100)/100)
347 GOSUB 1000
350 PRINT BB$, H
351 P=0
355 NEXT I
356 L=0
360 PRINT: PRINT
370 INPUT "BEGIN NEW TEST Y/N 44 BLACK SQUARE"; C$
375 IF LP=>TV THEN GOTO Y
380 IF C$="Y" THEN GOTO 31
390 END
400 LP=LP+1
410 Q$=Z$
420 PRINT "HEART": PRINT "TEST # SUBJECT": PRINT LP,Z$
430 IF LP>TV THEN GOTO 9
440 GOTO 300
1000 IF D(I)=254 THEN BB$="HORN": RETURN
1010 IF D(I)=253 THEN BB$="CLUTCH": RETURN
1020 IF D(I)=251 THEN BB$="BRAKE": RETURN
1030 IF D(I)=247 THEN BB$="THROTTLE": RETURN
1040 IF D(I)=239 THEN BB$="MIC": RETURN
1045 IF D(I)=500 THEN BB$="TEST CANCELLED": RETURN
1050 BB$=STR$(D(I)): RETURN
Parts List

1. ICA - 74LS05 - Inverter
2. 275-232 - Radio Shack - Relay
3. 2000 OHM Resistors

THIS SIDE FROM THE COMPUTER

+5 → ← +5

GNP → ← GND

PB1 → ← 1/6A

PB 2 → ← 1/6A

PA 0 → ← Brake Switch

PA 1 → ← Throttle

PA 2 → ← Clutch

PA 3 → ← Horn

Shutter Trigger

Not Used

Appendix B

Schematic of Electronic Interface

158
APPENDIX C
Schematic of Experimental Apparatus

Tape Recorder Instructions

Projection Screen

Timer

Handlebar Simulator

Microcomputer

Cathode Ray Tube

Slide Projector and Shutter
APPENDIX D
Letter, Number, Arrow and Landolt Ring Stimuli

Actual-Size Visual Stimuli
(i-r) Letter, Number, Arrow
Landolt Ring
Score Sheets:
An Experimental Analysis of the Static Visual Acuity of Novice, Experienced, and Formally Trained Motorcyclists.

<table>
<thead>
<tr>
<th>Correct response</th>
<th>Subject</th>
<th>Reaction</th>
<th>Time</th>
<th>Correct response</th>
<th>Subject</th>
<th>Reaction</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 left throttle</td>
<td></td>
<td></td>
<td></td>
<td>25 1 horn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 T clutch</td>
<td></td>
<td></td>
<td></td>
<td>26 1 horn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 left throttle</td>
<td></td>
<td></td>
<td></td>
<td>27 right throt.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 right brake</td>
<td></td>
<td></td>
<td></td>
<td>28 down brake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 left brake</td>
<td></td>
<td></td>
<td></td>
<td>29 down throt.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 L clutch</td>
<td></td>
<td></td>
<td></td>
<td>30 B clutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 up brake</td>
<td></td>
<td></td>
<td></td>
<td>31 right brake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 right throttle</td>
<td></td>
<td></td>
<td></td>
<td>32 M clutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 3 horn</td>
<td></td>
<td></td>
<td></td>
<td>33 up brake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 left brake</td>
<td></td>
<td></td>
<td></td>
<td>34 B clutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 up throttle</td>
<td></td>
<td></td>
<td></td>
<td>35 right brake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 left throttle</td>
<td></td>
<td></td>
<td></td>
<td>36 M clutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 7 horn</td>
<td></td>
<td></td>
<td></td>
<td>37 down throt.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 M clutch</td>
<td></td>
<td></td>
<td></td>
<td>38 9 horn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 T clutch</td>
<td></td>
<td></td>
<td></td>
<td>39 up throttle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 3 horn</td>
<td></td>
<td></td>
<td></td>
<td>40 L clutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 3 horn</td>
<td></td>
<td></td>
<td></td>
<td>41 7 horn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 L clutch</td>
<td></td>
<td></td>
<td></td>
<td>42 down brake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 T clutch</td>
<td></td>
<td></td>
<td></td>
<td>43 down throttl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 9 horn</td>
<td></td>
<td></td>
<td></td>
<td>44 B clutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 right throt.</td>
<td></td>
<td></td>
<td></td>
<td>45 up brake</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 up throttle</td>
<td></td>
<td></td>
<td></td>
<td>46 9 horn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 7 horn</td>
<td></td>
<td></td>
<td></td>
<td>47 1 horn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 left brake</td>
<td></td>
<td></td>
<td></td>
<td>48 down brake</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX F
Subject Consent Form (Ohio State Form HS-027)

THE OHIO STATE UNIVERSITY Protocol No. __________

CONSENT FOR PARTICIPATION IN SOCIAL AND BEHAVIORAL RESEARCH

I consent to participating in (or my child's participation in) research entitled:
An Analysis of the Static Visual Acuity of Novice, Experienced, Trained and Non-Motorcyclists During a Motor Task

______ Gary L. Winn (Principal Investigator) or his/her authorized representative has explained the purpose of the study, the procedures to be followed, and the expected duration of my (or my child's) participation. Possible benefits of the study have been described as have alternative procedures, if such procedures are applicable and available.

I acknowledge that I have had the opportunity to obtain additional information regarding the study and that any questions I have raised have been answered to my full satisfaction. Further, I understand that I am (my child is) free to withdraw consent at any time and to discontinue participation in the study without prejudice to me (my child). The information obtained from me (or my child) will remain confidential unless I specifically agree otherwise by placing my initials here ________.

Finally, I acknowledge that I have read and fully understand the consent form. I sign it freely and voluntarily. A copy has been given to me.

Date: ________________ Signed: ________________________

(Participant)

Signed: ______________________ Signed: ________________________

(Principal Investigator or his/her Authorized Representative) (Person Authorized to Consent for Participant - if Required)

Witness: ______________________

HS-027 (Rev. 10/91) — To be used only in connection with social and behavioral research.
LIST OF REFERENCES


