AN INVESTIGATION OF RISK HOMEOSTASIS IN A LABORATORY ENVIRONMENT

by

Rafael Marshall

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of Master of Science

in

Industrial and Systems Engineering

APPROVED:

Dennis L. Price, Ph.D.,
Co-Chairman

E. Scott Geller, Ph.D.,
Co-Chairman

Paul T. Kemmerling, M.S.

October 3, 1991

Blacksburg, Virginia

AN INVESTIGATION OF RISK HOMEOSTASIS IN A LABORATORY ENVIRONMENT

by

Rafael Marshall

Dr. Dennis L. Price, Co-Chairman

Dr. Scott Geller, Co-Chairman

Industrial and Systems Engineering

(ABSTRACT)

This study investigated whether risk compensation behavior would occur during a chemistry experiment due to the presence of protective equipment. This study also examined whether a homeostatic regulating mechanism exists for risk-taking behavior. Risk compensation and a homeostatic regulating mechanism for risk-taking behavior are both encompassed within the Risk Homeostasis Theory, which states that people accurately perceive and fully compensate for changes in risk.

Thirty-six subjects performed three trials of a short chemistry experiment either with protective equipment or without protective equipment during the first of two sessions. After the first session, half the subjects were required to switch from wearing protective equipment to not wearing protective equipment, or from not wearing protective equipment to wearing protective equipment. The time required to complete the task, the number of errors committed, and subtask measurement accuracy were tabulated.

Between-subject analyses did not reveal risk compensation behavior. Moreover, within-subject comparisons failed to show a significant risk compensation effect or the presence of a homeostatic regulating mechanism for risk-taking behavior. The results suggested that the Risk Homeostasis Theory may not explain sufficiently changes in behavior due to increases (or decreases) in perceived risk. The limitations of the present study were discussed. Suggestions and examples for research on different aspects of the Risk Homeostasis Theory were also provided.

ACKNOWLEDGEMENTS

I would like to thank Dr. Price, Dr. Geller, and professor Kemmerling for their time and encouragement throughout the many months it took to prepare this work. I would like to further thank Dr. Price for providing me with funds to conduct the research.

I would also like to extend a special thanks to Malay Dalal and Masae Tsukakoshi, whose advice, guidance, and friendship have been invaluable.

This thesis is dedicated to Celia Angeles and to my family.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This study investigated how the use of protective equipment affects the risk taking behavior of individuals during the performance of a laboratory task. The purpose of this study was to test aspects of Wilde's Risk Homeostasis Theory (1982b), which states that people accurately perceive and fully compensate for changes in objective risk. Wilde maintains that each person possesses a target level of risk under which he or she is willing to live. If the perceived risk for a certain activity were to somehow change, each person would alter his or her behavior in such a way as to bring the amount of risk back to its original level. Because individuals work to keep a certain risk level, the accident rate per time unit of that activity tends to stay constant.

According to the Risk Homeostasis Theory, the only way to reduce permanently the accident rate per unit time is through motivationally directed safety measures. Those measures not aimed at decreasing a person's desire to act in a less risky fashion will ultimately fail. Since the implementation of passive, nonmotivational safety measures are quite popular, Wilde's views stand in contrast to many programs currently taking place, such as the wide-scale use of driver-side air bags and energy absorbing materials in automobiles. If support is found for the theory, strategies for accident reduction may need to be reconsidered, with more thought going into the manipulation of the costs and benefits a person perceives when initiating risky behavior.

The Risk Homeostasis Theory has been forwarded as a general theory on human behavior in the face of risk (Wilde, 1982b; 1985), yet the majority of research generated by this theory has focused mainly on the risk taking tendencies of the road using population. This is not surprising, since the theory originated as an explanation of the limited reduction in traffic fatality rates following large-scale increases in safety belt use (Peltzman, 1975).

Nevertheless, research outside this realm would not only provide information on the generalizability of the theory, but may also provide an experimental situation more suitable for the observation and evaluation of risk compensation behavior. The present study performed all of these functions. It tested aspects of Risk Homeostasis in a different environment, allowing the researcher to make detailed observations of subject performance during a chemistry experiment.

Objectives

This study investigated the presence of a risk homeostatic mechanism and of risk compensation behavior in male college students as they performed a chemistry experiment under different degrees of perceived risk. The study was conducted to assess the validity of the Theory of Risk Homeostasis as formulated by Wilde (1982a; 1982b). A chemistry setting was selected because of the ease in which the perceived risk of the situation can be manipulated without significantly affecting the objective risk. Additionally, industrial laboratory environments contain many hazards which could easily result in accidents, making data on risk taking behavior useful for real-world application.

LITERATURE REVIEW

Background

The last 20 years have seen increased concern toward the reduction of automobile accidents and fatalities. Seat belt use laws in many states and the pressure for mandatory installation of airbags exemplify this concern. Yet, a few researchers have questioned whether effort has been misguided towards measures that produce only minimum effects, if any at all. Adams (1982, from Evans, 1985, p. 560) compared the incidence of fatalities in 13 countries which had mandatory seat belt use laws to four countries which had no such law. He concluded that the law did not reduce fatalities. Herms (1972) examined the effects of pedestrian crosswalks on pedestrian safety and found that, instead of producing positive effects, the incidence of fatalities actually increased significantly. On the other hand, the imposition of the 55 mile per hour speed limit in 1974 reduced total traffic accidents by impressive amounts (Evans, 1986), even on roads not affected by the limit.

Traffic Risk Models. Various theories have been proposed to explain why accidents occur and which safety measures have the best chance of increasing traffic safety. Among these are the Zero Risk Model developed by Naatenen and Summala (1976), Fuller's Threat Avoidance Model (1986), and Van der Molen and Botticher's Hierarchial Risk Model (1986).

Naatenen and Summala (1976; 1988) stated that a driver drives in a way which satisfies prevailing motives, but adopts a safety margin which causes him/her to feel no subjective risk. If risk is felt, the driver attempts to alleviate this feeling through immediate behavioral changes. It is the tendency of the driver to avoid choices that may result in noticeable increases in risk. Although risk is generally avoided, the driver's subjective risk threshold is distorted by learned expectancies, adaptation to current risk situations, and extra motives (e.g., in a hurry, trying to impress friends). This results in an increase in the subjective risk threshold and in the adoption of a narrowed safety margin. The narrowed safety margin may not provide the driver with sufficient reaction time in the event of a high risk situation, thereby increasing the probability of an accident.

Fuller's Threat Avoidance Model (1984) postulated that when confronted with a potentially aversive stimulus, drivers modify their driving behavior in such a way as to avoid the stimulus while maintaining the rewards of the present behavior as much as possible. The potentially aversive stimulus is sensed through perceptions of speed, the road environment of the intended path, and the drivers' ongoing capabilities (i.e., skills and experience). Once the situation has been assessed, drivers either act in anticipation of the threat or they delay response until the threat is realized. In the case of the latter situation, the threat may be illusory; but if it is real, a delayed avoidance response must be performed. Failure to respond quickly and correctly may result in an accident.

Van der Molen and Botticher's Hierarchical Risk Model (1988) postulated that the driver's thought process proceeded through three hierarchical levels: the Strategic, Tactical, and Operational levels. At the Strategic level, the highest in the model, route planning, trip duration estimations, desired cruise speed, and other broad plans are made. At the Tactical level, plans are more concrete, such as deciding to change lanes, increasing speed to overtake a truck, or braking for an upcoming stoplight. The plans made at the Tactical level are carried out at the Operational level. At this level, basic skills and movements are performed to satisfy the decisions made at the higher levels, such as pressing on the accelerator, steering the wheel, and looking in the rear-view mirror. Also carried out at the Operational level are emergency decisions stemming from direct feedback from the external environment. The plans made at all levels are a function of the physical environment, the driver's existing internal representations, motivation, expectancies, and feedback from prior actions and plans.

Van der Molen and Botticher divided the judgments made during the Strategic and Tactical levels into those that involve risk and other judgments. Since the motivations and expectations that form judgments can change from one situation to another, the authors stated that the level of risk a driver is willing to take or accept also changes with the situation. Yet, while the accepted risk might change at the Tactical level (e.g., amount of overtaking maneuvers to be performed), Strategic level risk may stay constant (e.g., cruise speed to be maintained).

Both Fuller's (1984) and Van der Molen and Botticher's (1988) theories have been influenced by Wilde's Risk Homeostasis Theory (1982b). Formally outlined in 1982, this theory discounted many recent safety measures as only temporary solutions to accident reduction. It posited that only those measures directed at motivational risk reduction will have long-term effects on traffic safety. Because of this somewhat bold assertion, as well as what many consider ambiguous language which has allowed Wilde to modify its basic tenets to discount seemingly damaging data, this theory has been a major source of controversy for much of the last decade.

Risk Homeostasis Theory

Overview. According to the Risk Homeostasis Theory (RHT), "traffic accident rate per time unit of road user exposure is the output of a closed loop regulatory process in which the target level of risk operates as the only controlling variable" (Wilde, 1986a, p.379). As illustrated in Figure 1, each road user compares the perceived risk level of the present situation with the target level of risk he/she is willing to accept, and then adjusts behavior accordingly to bring the two into alignment. Whether the subsequent behavior

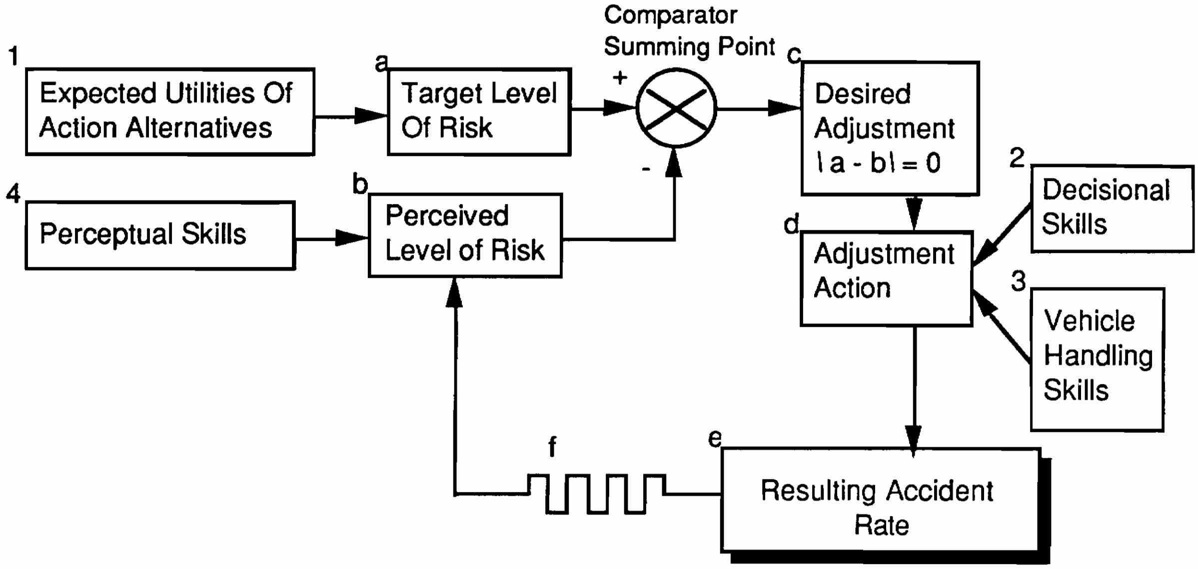

Figure 1: Homeostatic model relating accident rate to driver behavior (Wilde, 1982b).

Note: box 'c' reflects changes made in Wilde (1986).

has the desired result depends on the individual's perceptual, decisional, and executional skills. The comparison process is thought to occur mainly at the "pre-attentive level" (Wilde, 1988), but can be brought into conscious awareness by simply questioning the user, or through stimulus events beyond the just noticeable difference in the subjective estimation of danger (Wilde, 1982b).

The target level accepted by the user is determined by the subjective maximization of benefits to costs involved (Wilde, 1986a). The cost/benefit ratio is dependent on long-term, short-term, and momentary fluctuations in each road user's motivational state, which in turn is a function of economic, cultural, and person-related variables. Although short-term and momentary motivational factors may cause the target level to fluctuate, these variations averaged over time will equal the accepted target level if road characteristics are kept constant.

The target level of risk can vary widely between individuals. Wilde (1982a) interpreted this as essential for meeting the functional requirements of society. Society requires that risk takers fill positions such as law enforcement officials, firefighters, and coal miners, and risk avoiders take positions such as safety inspectors and child-care providers. The setting of a target level carries with it a certain likelihood of objective accident risk. For a given population, the sum of the product of each accident multiplied by its severity rating gives the accident rate by which the population is willing to live to obtain the benefits that correspond to the target level of risk accepted. According to Wilde, the cost (severity) of an accident can range from time and productivity loss, to property damage, to injury and death (Wilde, 1988).

To reduce the accident rate, it is necessary to reduce people's target level of risk. According to Wilde (1982a), nonmotivational safety measures will not affect the target level. As shown in Figure 2, such a measure would only work inside the feedback loop.

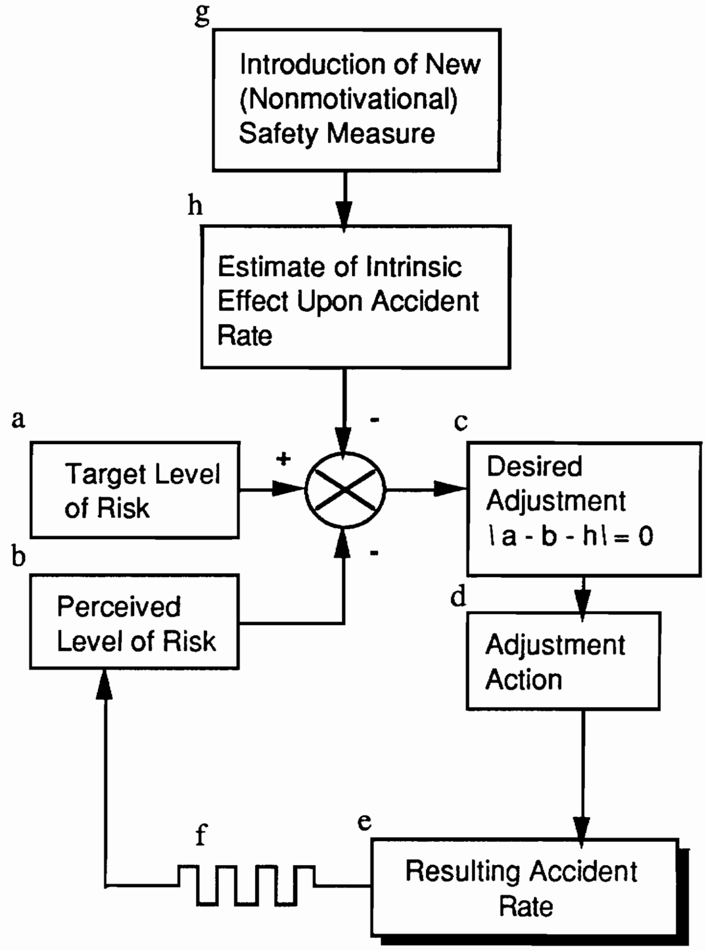

Figure 2: Model of effect of introducing a new, nonmotivational, traffic safety measure (from Wilde, 1982a). Note: box 'c' reflects implied changes from Wilde (1986).

Individual drivers would assess the intrinsic effects of the measure through experience and other external cues, and adjust behavior in a way that would eventually synchronize the new risk factor with the target level chosen. The time lag for resynchronization is not fully specified, but Wilde does volunteer not more than two years as a reasonable range (Wilde, 1982a).

For example, if all convertible owners were suddenly required by law to wear crash helmets while driving, RHT posits that these drivers would compensate for this added protection in such a way which would eventually bring this population's accident rate per unit time back to its original level. The denominator specified in the last statement is very important in understanding Wilde's theory, for were these drivers to compensate by speeding, this may reduce the accident rate per kilometers travelled, but not per time spent on the road (assuming that the drivers also drove more often). Therefore, the theory does not state that non-motivational safety measures are useless, only that there are better ways of reducing the accident rate.

According to Wilde, motivational safety measures are the only effective solution for reducing the target level of risk. As was mentioned previously, the target level accepted by the user is determined by the subjective maximization of the cost/benefit ratio. Wilde identified four controlling factors for this ratio:

1) the perceived benefits of cautious behavior

2) the perceived costs of risky behavior

3) the perceived benefits of risky behavior

4) the perceived costs of cautious behavior

By increasing either of the first two or decreasing either of the last two, the target level of risk can be reduced. For example, Wilde stated that the costs of cautious behavior can be lessened by introducing flextime into the work schedule; the costs of risky behavior can be increased by larger fines for failing to use vehicle seat belts; the benefits of risky behavior can be reduced by paying taxi drivers per unit time and not distance; and the benefits of cautiousness can be increased by offering rewards for accident-free driving (Wilde & Murdoch, 1982, p 882).

To support his theory, Wilde relied mainly on studies performed by Peltzman (1975), Adams (1981), Taylor (1964), and Partyka (1984). The Peltzman and Taylor studies were also the basis for much of the hypotheses that compose Wilde's theory. Using a time-series model, Peltzman (1975) studied the effectiveness of the newly introduced vehicle safety standards (i.e., shoulder and seat belts, energy-absorbing steering column, penetration-resistant windshield, padded instrument panel, and dual braking system) in the prevention of highway deaths. He concluded that these safety features triggered an increase in "driving intensity", resulting in no significant deviation from the pre-standard highway deaths for the road-using population. Here, the road-using population refers not only to drivers and their occupants, but also to pedestrians, motorcyclists, bicyclists, and the like. Therefore, while traffic safety features reduced the number of automobile occupant fatalities, this decrease was offset by an increase in deaths in the other subgroups.

Adams (1981, from Wilde, 1984) compared the number of fatalities in 13 countries having seat belt use laws to four countries that did not. He found that, although the countries with a seat belt use law did record a drop in the average number of fatalities, the no-law countries recorded a drop greater than that of the law-countries. Wilde (1984) interpreted this as supporting RHT, since a large number of the drivers in the seat belt use law countries may have initially overestimated the safety factor seat belts afforded, thereby causing the number of fatalities to rise with respect to the no-law countries. Interestingly, Wilde made no mention why the average overall number of deaths were falling in the first place, a circumstance that cannot be explained by RHT.

Wilde used Taylor's (1964) pioneering work on risk compensation to support the idea that drivers are, on the average, good judges of situational risk. Taylor measured galvanic skin response in a sample of drivers over 40 different driving environments and found the level of GSR activity generally did not fluctuate relative to the road condition encountered. Yet, the mean GSR rates measured while driving were approximately 50 times greater than those measured when the subjects were quietly sitting in the laboratory. Taylor interpreted this to mean that drivers adjusted their driving behavior to keep the perceived risk at a constant level per time-unit of driving.

Wilde often cited Partyka's study (1984) as showing the effect of economic trends on the degree of risk taking in the population. Adjusting for the 1974 oil embargo and the implementation of the 55 mile per hour speed limit law, Partyka found that traffic fatality trends from 1960 to 1982 were quite consistent with changes in unemployment figures. As the number of unemployed workers increased, the incidence of traffic fatalities decreased by a fixed magnitude. Inversely, the number of traffic fatalities tended to rise as the number of employed workers increased. Partyka noted that the primary intent of the model was not to imply a causal relationship between employment and traffic accidents, but to point to coinciding changes among variables. Nevertheless, Wilde (1984) interpreted the strong positive relationship as verification of how economic trends can effect traffic accident rate.

Recently, Wilde (1991) suggested a way to link incentives with economic trends to bring traffic accident rates down annually. By nationalizing auto-insurance, the government could reduce the annual traffic accident toll, even during times of economic upswing, by increasing insurance surcharges as a penalty for being in an accident, and offering discounts and rewards for accident free driving. Wilde stated that such a program would probably not be carried out satisfactorily by private insurance industries because a dramatic reduction of accidents would ultimately be against their money making interests.

Points of Controversy

Since its formal introduction in 1982, the Risk Homeostasis Theory has generated much debate over its premises. Table 1 provides a list of articles supporting or disputing the theory. Two critical articles, and their respective rejoinders, are covered in some detail below to elaborate on other important aspects of the theory and to expose its weaknesses.

One of the theory's staunchest critics is McKenna (1985), whose major concern lay with the assertion that classical safety measures generally have no lasting effect. He saw four underlying assumptions as crucial to Wilde's position (McKenna, 1985, p. 489):

1) People have a simple straightforward representation of accident risk.
2) People can detect all changes in this accident risk.
3) People can, over time, completely compensate for these changes.
4) People cannot be discouraged or prevented from compensating for changes in accident risk.

With regard to the first, McKenna argued that how information is presented can greatly impact risk perception. Slovic, Fischhoff, and Lichtenstein (1978) found that presenting subjects with the probability of a traffic accident per lifetime (a probability of 33%) resulted in more positive attitudes towards safety belts than when presenting the probability of an accident per trip (a probability of $3 \times 10^{-5}\%$, or one chance in 3.5 million). McKenna questioned the ability of people to comprehend and compensate accurately for a 1 in 3.5 million chance of being involved in a fatal accident in the daily business of driving.

TABLE 1

Summary of Articles Written in Support/Contention with the Risk Homeostasis Theory (RHT)

Study	Principle Opinions and Findings
Wilde, G.J. (1982a)	• Formally introduced the Risk Homeostasis Theory
Slovic, P. and Fischhoff, B. (1982)	• Questioned how user habits, adaptation, and "invisible safety measures" fit into RHT
Graham, J. D. (1982)	• Stated that RHT was too narrow in scope. Risk was only another commodity to consider along with safety, pleasure, etc....
Cole, G. A. and Withey, S. B. (1982)	• Stated that Wilde blurred distinction between individual and aggregate
Wilde, G. J. (1982b)	• Rejoinder to Slovic and Fischhoff (1982), Graham (1982), and Cole and Withey (1982)
Wilde, G. J., and Murdoch, P. A. (1982)	• Elaborated on steps towards long-term accident reduction
McKenna, F. P. (1982)	• Cited studies against RHT and questioned how people were able to estimate accurately extremely small probabilities
Wilde, G. J. (1984)	• Rejoinder to McKenna (1982). Drew attention to the importance of the correct denominator (accidents per time unit)
McKenna, F. P. (1985)	• Examined what he believed to be the critical assumptions made by RHT
Wilde, G. J. (1985)	• Rejoinder to McKenna (1985). Corrected McKenna's suppositions concerning RHT
McKenna, F. P. (1985a)	• Criticized Wilde's selective use of methodological criteria
Wilde, G. J. et al. (1985)	• Created study to test RHT. Found partial support for theory
Evans, L. (1985)	• Expressed skepticism that individuals have an all-pervasive desire to maintain risk at a constant level.
Evans, L. (1986)	• Examined traffic accident data and found all to be incompatible with RHT
Wilde, G. J. (1986)	• Rejoinder to Evans (1986). Argued that Evans did not take into account relevant economic changes and made various statistical errors
Evans, L. (1986a)	• Commented on the lack of solid support for RHT

TABLE 1 (Continued)

Summary of Articles Written in Support/Contention with the Risk Homeostasis Theory (RHT)

Study	Principle Opinions and Findings
Wilde, G. J. (1986a)	• Pointed to applicability of RHT for future research and application in the areas of road accidents and other accident domains
Shannon, H. S. (1986)	• Analyzed British road accident data and found no support for RHT
Adams (1988)	• Discussed the untestability of RHT and concluded that the theory may only be good for generating thought
Janssen, W. and Tenkink, E. (1988)	• Argued that risk homeostasis rarely occurred and that risk should be considered only as part of a more general utility maximization process
Wilde, G. J. (1988)	• divided RHT into 15 components and discussed evidence in opposition and support of the theory
McKenna, F. P. (1988)	• Noted several methodological and theoretical inconsistencies in RHT and presented studies that ran counter to the predictions of RHT
Streff, F. M. and Geller, E. S. (1988)	• Found support for risk compensation mainly through within-subject comparisons. Findings suggested that the occurrence of risk compensation was dependent on individuals being able to compare one situation to another
Wilde (1991)	• Suggested that traffic crashes could be reduced by nationalizing auto-insurance

For the presumed ability of the road user population to detect all changes in risk level, McKenna reintroduced a question originally expressed by Slovic and Fischhoff (1982) who had voiced the possibility of a safety measure being psychologically invisible to most road users. Wilde had responded by directing attention to the feedback loop, stating that, with time, the road using population will notice the change in risk level and eventually compensate for it. McKenna expressed dissatisfaction with this explanation, on the grounds that a time lag was not strictly specified in the theory. Unless better defined, any study examining RHT could be discounted for not allowing enough time to expire for a change in risk to be perceived and acted upon by the road using population. McKenna added that although he agreed there are many factors operating on the road user, it is doubtful that maintaining a constant level of deaths is one of their major concerns. He concluded by pointing again to the difficulties of risk assessment, stating, "How would the ordinary driver know the oncoming lamp post or other piece of roadside furniture had energy absorbing characteristics?" (McKenna, 1985, p. 491).

Wilde often cited Peltzman (1975) for evidence of complete compensation towards safety directed measures. As previously mentioned, Peltzman concluded that the introduction of safety measures caused drivers to compensate by increasing their driving intensity, resulting in no change in the accident rate for the road using population. He concluded that the decrease in driver risk offered by the safety measures was offset by an increase in pedestrian fatalities, revealing a shift in risk from driver to pedestrian. McKenna contended that this shift in risk is incompatible with RHT. Although the aggregate risk may have stayed constant, individual risk levels did not. McKenna questioned why it is that pedestrians were willing to accept this higher risk instead of compensating to bring it to previous levels. McKenna also noted that Peltzman's analysis has been under fire, citing critical studies by Joksch (1976) and Robertson (1977) which argue that some measures used in the multiple regression models were theoretically inappropriate or used in a distorted fashion, and that other likely influential factors were ignored.

Lastly, McKenna directed attention to the effectiveness of lower speed limits in reducing the accident rate. He also cited the successful use of visual illusions, specially marked pedestrian crosswalks, and invisible safety measures as evidence that people can be prevented or discouraged from compensating fully.

In a rejoinder to McKenna's article, Wilde (1985) stated that RHT does not assume every individual has an accurate picture of the risks involved -- this is frankly why accidents happen. But when averaged across all road users, the population's representation of risk, observed through the accident rate per capita, is quite constant.

To the arbitrariness of the time lag, Wilde (1985) stated that the time lag was linked to the definition of homeostasis, but had yet to be measured by empirical investigation. From studies done previously, the lag would not seem to be greater than two years, with some being so short as to be hard to detect reliably. If a safety measure produced a change in the accident rate below the just noticeable level, it may have a lasting effect, but since the effect of safety measures is additive, the risk reduction effects will eventually be noticed and compensated for until it is again below the just noticeable difference. The present author would note, however, that if the effect of a safety measure was initially below the just noticeable difference, effectively making it invisible, it is not clear how the average road user would know what compensatory behavior to take once the additive nature of several safety measures cause the threshold to be exceeded and the reduction of accidents to be eventually noticed. It is illogical to assume that the nature of all these previously invisible measures would immediately be known to the road user once threshold is exceeded. This is of importance because it would seem likely that drivers would compensate differently in response to improvements in automobile braking systems as opposed to improvements in energy absorbing light posts.

In response to the risk shift from drivers to pedestrians, Wilde wrote that RHT considers the road user population collectively. He suggested that sub-populations not directly addressed by the safety measure may be slower in becoming aware of changes, consequently resulting in a delayed reaction time. Wilde later expanded on this idea by adding that a "long-term shift in the proportional contribution of the accident loss in a particular mode to the total per capita traffic accident loss will occur if the intervention leads people to spend more of their road-travel time in one mode rather than in another" (Wilde, 1988, p.448). In other words, if the cost/benefit ratio for driving an automobile is made more attractive by a safety measure, more people will travel by automobile than by some other mode of road transportation (e.g., bicycle, motorcycles, walking). This shift would increase the proportion of automobile accidents, while decreasing the amount of traffic loss for the non-addressed modes -- both, theoretically, to their previous levels.

As for McKenna's last assertion, Wilde (1985) noted that it is one of the basic tenets of RHT that accidents per capita can be brought down. Nonmotivational measures could be effective in reducing accidents per kilometer travelled, and such implementations would allow users to "enjoy greater mobility per time unit in return for the same accident loss" (Wilde, 1988). However, motivationally based safety measures are the only way to reduce the accident rate per time unit and per capita. He concluded by stating that the "proponents of RHT are no more pessimistic... than physicians who tell their patients that pneumonia cannot be cured with aspirin, while at the same time handing over a prescription for antibiotics" (Wilde, 1985, p. 1536).

In another commentary of Wilde's theory, Evans (1986) attempted to "definitively answer" whether there was any validity to risk homeostasis. To do so, Evans examined seven traffic-related variables for concordance with RHT predictions:

1) long-term trends in traffic accident rates.

2) short-term trends in traffic accident rates following a one time perturbation.

3) long-term trends of traffic accident rates after a major perturbation.

4) traffic accident rates on different types of roads.

5) response of traffic accident rates to changes in laws mandating the use of safety equipment.

6) the traffic accident rate of individuals over time.

7) long-term trends in overall accident rates (all causes).

For all seven cases, Evans found that the statistics contradicted the predictions that would have evolved from RHT. For example, while the theory predicts that accident rate per unit time of driver exposure is independent of road geometry (Wilde, 1982b), Evans cited data from the Department of Transportation showing a distinct difference between fatalities per billion hours on interstate highways and secondary roads that received federal aid.

Evans referred to data from an "almost perfect natural experiment" to refute RHT's prediction that laws mandating the use of safety equipment will have no effect on the number of fatalities. During the mid-1970's, 26 of the 50 states decided to repeal a law requiring motorcyclists to wear helmets. On the average, the states that had repealed the law experienced a collective increase of 28% in motorcyclist deaths when compared to those states that kept the law intact. Evans concluded by explaining that "the risk homeostasis theory suffers from two major deficiencies: (1) no evidence supporting it and (2) much evidence refuting it." (Evans, 1986, p. 93). Additionally, Evans stated that the theory is ambiguous in wording and confusing in what may, at face value, seem like simple concepts.

Wilde (1986b) responded to Evans' seven variables by pointing to statistical errors, restrictive use of data, narrow or erroneous understanding of the theory, and questionable interpretation of data. In response to the data used to refute the idea that risk is independent of road geometry, Wilde again argued that Evans did not take into consideration the whole population of road users. As for the declining accident rate, Wilde attributed this to the energy crisis of the mid-1970's, which greatly affected cost/benefit maximization levels, and therefore target risk levels.

Towards Evans' data on the repeal of the motorcycle helmet law, Wilde cited the absence of random assignment, suggesting the factors causing some states to repeal the law may have also influenced the subsequent number of fatalities. In addition, although the fatalities were considered as a function of population growth and geography, seasonal variations were not taken into account, thus ignoring a factor of significance to motorcyclist death rates.

In a subsequent commentary on Wilde's remarks, Evans dismissed most of them as weak and "ad hoc" in nature (Evans, 1986a, p. 103). For the most part, the present author agrees. Most of Wilde's complaints towards the motorcycle helmet study were of minor significance to the general results presented by Evans. Additionally, Wilde placed much emphasis on the "energy crisis" for the reduced risk-taking appetite of road users during the mid-seventies, but did not explain sufficiently, in terms of his theory, why the accident rate has not returned to previous levels after the "energy crisis".

Testability

Most of the studies cited by the critics and supporters of RHT were not originally formulated to test the theory. This is evident in the articles of McKenna and Evans, as well as in the rejoinders of Wilde. Because of this, the number of interpretations have tended to equal the number of researchers addressing the issue. The Risk Homeostasis Theory has been in the literature for over a decade, yet relatively few studies have been published either in support or in contention with its premises. This lack of data seems due to four factors:

1) the possibility of risk transfer into other domains of activity

2) the difficulties in accurately testing risk in a laboratory environment

3) the large number of variables that influence risk taking

4) ambiguity as to the unit of analysis (individual vs. aggregate)

As previously mentioned, Wilde (1986a) incorporated the possibility of risk transfer from one subpopulation of road users to another into his theory. This in itself presents difficulties in testing, but Wilde (1982b; 1986a; 1988) also mentioned the possibility of risk transfer from one domain of general activity to another (e.g., a decrease in smoking related deaths offset by deaths due to alcohol consumption, traffic accidents or other causes). Therefore, it would be hypothetically possible to refute findings which show that safety belt use was successful in reducing the accident rate by pointing to an increase in drug-related deaths across the population (Wilde appeared to be doing just that in Section 10 of his 1986a article). It should be noted that such an explanation would not answer whether the same people who benefitted from safety belts were the same ones who were taking more risks with drugs, which would have to be the case for individuals to maintain their target level of risk. For the purpose of this experiment, it was assumed that such shifts in risk do not occur. This would be compatible with RHT at the individual level, where target risk must be maintained.

A difficulty related to the problem mentioned above lies in deciphering the level at which RHT applies -- the individual or the aggregate level. The RHT model relies on the perceptions and behavioral actions of the individual as the mode of illustration. The model then shifts to the aggregate in order to explain time lag and risk transfer, to name a few. McKenna (1985b; 1985a; 1988) contended that explanations used to interpret aggregate data are often in conflict with RHT assumptions that modifications in behavior will always occur, will always be compensatory, and will always be complete when dealing with nonmotivational safety factors.

Wilde has repeatedly stated that the proper unit of analysis for RHT as a whole is the aggregate (Wilde, 1982b; 1985; 1986a; 1988). Yet, the individual is clearly important when explaining the underlying premises of the theory. Both Cole and Withey (1982) and McKenna (1985a) have stated that any data revealing a shift in aggregate behavior will always draw attention to the individual actions(s) that were instrumental in bringing about the change. It is from this knowledge that safety measures can be examined for effectiveness.

In a notable study that examined both within- and between-subject responses to risk, Streff and Geller (1988) required subjects to drive a 5-hp go-kart around a clay track, with or without the use of safety belts, in order to observe any evidence of risk compensation. Subjects were assigned to one of four conditions:

1) using safety belts during both sessions of the experiment

2) not using safety belts during either of the two sessions

3) using safety belts only during the first session

4) using safety belts only during the second session

Little support for risk compensation was found between the four conditions. However, subjects in the last two conditions reported changes in perceived safety, with those in the third condition feeling less safe in the second session, and those in the fourth feeling more safe. It was also found that subjects who went from not using seat belts in the first session to using seat belts in the second session (fourth condition) increased their driving speed during the second phase significantly more than subjects who used the safety belt for both phases. Therefore, it appeared as though individuals needed a basis of comparison (safe vs. unsafe) before compensation behavior could occur. Between-subject studies would not be able to measure risk compensation, since individuals would not have had an opportunity to compare one situation with another.

Wilde has expressed strong reservations towards the ability of risk to be tested accurately and validly in a laboratory environment, stating that "resorting to laboratory and simulation studies may be methodologically pleasing (and morally innocuous), but it is doubtful that the theory in question can ever be cogently tested under such contrived conditions" (Wilde, 1982b, p. 251). Regardless of this view, Wilde himself attempted to support RHT through a computer game experiment (Wilde, Claxton-Oldfield, & Platenius, 1985), finding it possible to study at least some aspects of the theory in a non-safety-related fashion. In this study, male participants cancelled a randomly lit stimulus light when they felt that 800 milliseconds (ms) had elapsed. The closer a subject came to 800 ms, the greater the financial reward they received. If 800 ms was exceeded, a penalty of one of two predetermined probabilities (.3 or .7) was incurred. Before taking part in the video game, subjects were assessed for risk taking tendency using Zuckerman's Stimulation Seeking Scale (1971). It was expected that while all subjects would finish with roughly similar amounts of money (homeostasis), the "risk takers" would earn their money by risking the possibility of exceeding 800 ms in order to receive larger financial rewards, while "risk avoiders" would earn their money by making safer, though less rewarding responses. It was also hypothesized that when the probability of loss was increased from

30% to 70%, the subjects would respond in a less risky fashion, leading to a lower average response time. The results showed that, although risk avoiders did tend to take more chances than did the risk takers, the difference was not statistically significant. It was also found that the mean response latency did decrease from 682 ms on the 30% probability of loss to 653 ms on the 70% probability of loss, in accordance with the predictions. From the results, Wilde et al. (1985) concluded that although behavioral compensation did occur, complete homeostasis was not observed.

Unfortunately, Wilde's reservations as to the ability of risk theory to be tested properly in a safety-related field may be well founded. Under Wilde's theory, homeostasis exists in terms of losses (accidents, fatalities) per unit time. Feedback, by which judgments of risk level are made, is also a function of the losses experienced directly or indirectly by a certain population over an extended period of time. Obviously, a study contrived to test risk homeostasis in such a way would be morally and ethically infeasible. Although it would be very difficult to test every aspect of the Risk Homeostasis Theory as related to safety, mechanisms by which the theory is said to occur may be tested as part of a risk compensation study.

Since people make their judgments of risk from their own day-to-day experiences and from news received from newspapers, television, and friends (Wilde, 1982b), it would follow that the amount of risk a subject initially perceives may not necessarily be the amount of objective risk present. The present study controlled the perceived risk of the experimental situation while minimizing the objective risk.

Testing Environment

The Risk Homeostasis Theory was originally introduced as an attempt to explain the occurrence of traffic accidents. Yet, Wilde has on many occasions expressed the possibility of expanding the applicability of the theory beyond traffic safety (Wilde, 1982b; 1986a; 1988), stating that "...it would be surprising indeed, if mechanisms that control people's risk-taking on the road were fundamentally different from those in other behavior areas, like occupational life, health habits, sports, and so forth." (Wilde, 1982b, p. 222). The present study investigated whether aspects of RHT were truly applicable to safety-related environments other than the road.

Briefly, the present study tested aspects of RHT in an environment which roughly simulated that of a chemistry laboratory. Knowledge of risk-compensation behavior in this environment could be invaluable at a time when highly dangerous lab samples are handled frequently. Subjects were asked to perform six trials of a chemistry experiment over two sessions, as accurately and as safely as possible. Each subject in the experimental groups was randomly assigned protective equipment on either the first or second session. Subjects in the control groups were either given protective equipment or not given protective equipment during both sessions. The directions, warnings, and chemicals were identical, whether or not protective equipment was provided. Unknown to the subjects, the chemicals used were very safe.

Feedback on risk/safety, an important part of Wilde's theory, was provided through experimenter comments and actions. Whether a subject performed with or without protective equipment, the experimenter always wore goggles, an apron, and surgical gloves and always acted in a cautious manner. Additionally, the experimenter required that each subject wash their hands at the end of the task if any of the chemicals came in contact with their skin. Lastly, subjects were asked to read, and were read, a list of safety measures that had to be followed during the task.

It was expected that subjects who performed the task without protective equipment would act in a more cautious manner than subjects with protective equipment. Cautious behavior in the part of those without protective equipment was expected to result in fewer mistakes the need for more time to complete the task. Accuracy was also expected to increase for those without protective equipment due to the higher degree of attention and caution required in preventing an accident from occurring. Subjects with protective equipment were expected to act in a less cautious manner and pay less attention since the perceived cost of an accident is lessened by the use of the equipment.

"Accident rate", as defined for the Risk Homeostasis Theory, is the product of the frequency and the severity of an accident, totalled over a given time span. A population's target risk level is measured through its accident rate. For this study, "accidents" were defined as procedural errors made during the task. Because objective risk in this study was almost nonexistent, severity did not enter into the equation and was dropped from the equation. Since Wilde's theory predicts that a nonmotivational change in risk-level will not change a population's accident rate, it was expected that the number of errors per minute for each group in this study would be constant.

Hypotheses

Hypothesis 1 predicted between-subject effects during the first session among the four conditions. Hypotheses 2 and 3 predicted within-subject effects during the second session. Finally, Hypothesis 4 investigated whether a homeostatic regulatory mechanism truly existed, as stated in Wilde's theory.

1) Subjects performing without protective equipment in the first session will perform in a less risky manner (take more time, make less errors, be more accurate) during the first session than subjects with protective equipment.
2) Subjects who switch from having protective equipment in the first session to not having protective equipment in the second session will perform in a less risky manner (take

more time, make less errors, be more accurate) during the second session compared to those performing with protective equipment throughout both sessions.

3) Subjects who switch from not having protective equipment in the first session to having protective equipment in the second session will perform in a more risky manner (take less time, make more errors, be less accurate) during the second session compared to those performing without protective equipment throughout both sessions.
4) Error rate (number of errors per task, divided by task completion time) will be constant across all four conditions for both sessions.

METHOD

Subjects

Thirty-six male subjects between the ages of 18-25 were recruited from the Virginia Polytechnic Institute and State University (Virginia Tech) campus for the study. All were required to sign an informed consent form prior to their participation (Appendix A). Payment of $15.00 was given for taking part in the study.

Apparatus

Subjects were asked to perform a freshman-level acid-base titration task in an actual laboratory setting. The chemicals used for this experiment consisted of phenolphthalein solution to indicate acid neutralization, 0.1M of citric acid, and 0.2M of ammonia. A stopwatch was used by the researcher to time each trial.

Citric acid was chosen as the acid for several reasons. First, at a concentration of 0.1M, its pH value was 2.09. This meant that it had the acidity level of lemon juice. Second, citric acid had high name-recognition. It was determined that this point would be important during subject debriefing. Third, the acid was also quite inexpensive to obtain. The base, 0.2M of ammonia, had a pH value of approximately 11.27. This pH value was much less than that of household ammonia. There was an eye-wash within 50 ft. of the work area in the event of eye contact with the acid or base. A first aid kit was present in the laboratory. Additionally, Kelly Matthews, a Ph.D candidate in the Chemistry Department, assisted in preparing the chemistry equipment and solutions so as to minimize hazards. The objective risks involved in this study were, therefore, minimal.

The chemistry apparatus were kept simple to allow subjects to learn their appearance and purpose within a few minutes. A list of the apparatus is provided in

Appendix C, along with the laboratory directions and procedures. Brief descriptions of a few of the more important items are also provided.

General Procedure

First session. Each subject was randomly assigned to one of four treatment conditions. Subjects in the first treatment condition were provided with goggles, surgical gloves, and aprons during the first session, but were not provided with protective equipment during the second session. Subjects in the second treatment condition were not given goggles, surgical gloves, and apron during the first session, but were provided with the protective equipment during the second session. Subjects in the third treatment condition were given protective equipment during both sessions, and subjects in the fourth condition were not given any protective equipment during both sessions.

Each subject was tested individually. Once the necessary forms were read and signed by the subject, the researcher spent 15 minutes acquainting the subject with the equipment and procedures of the titration task. The researcher demonstrated the titration procedure and allowed the subject to practice the task using tap water. Once the subject was familiar with the equipment and procedures, general information and precautions pertaining to the experiment were read. The researcher suggested a time-for-completion of nine minutes for the task. There were no penalties given for exceeding the suggested completion time, but the subject was asked to use this time to pace his progress. The subject was then told to read the instructions carefully and completely before starting the actual task in order to budget time. Since questions during the task were not allowed, the subject was told to voice any concerns prior to initiating the task. The goggles, surgical gloves, and apron were then provided if the subject had been assigned to receive them.

Each subject was given a prepared data sheet (Appendix B) to record titration measurements. The researcher was in the same room with the subject, stationed approximately 15 ft away. This served the dual purpose of safety and ease of data collection. All trials were also videotaped to ensure accurate data collection. Time started when the subject started the timer provided, indicating task initiation. The subject was not allowed to begin the task until the timer had been pressed (timer emitted a "beep" once activated).

The titration task itself contained five subtasks. A breakdown of the five subtasks are given in Table 2. For the first subtask, the subject was instructed to measure accurately 5 ml of a solution, named "S", into a pre-weighed graduated cylinder. Solution "S" was the indicator. Its identity was kept from the subject to aid in the perception of risk. Once this had been done, the subject was instructed to place the graduated cylinder containing solution "S" on the digital scale provided and record the weight (in grams) displayed. The accuracy by which the solution was measured into the graduated cylinder was obtained from the combined weight of the solution and cylinder. The subject then placed the solution into a specified erlenmeyer flask. During the second subtask, the subject was instructed to weigh 0.5 grams of the acid onto a piece of weighing paper, record the weight obtained, and add the acid to the flask containing solution "S". The third subtask consisted of the subject placing 15 ml of water and two drops of indicator into the flask. The amount of water and drops used were recorded. The fourth subtask required the subject to fill a specified buret with the base until the 40 ml line. The fill point was recorded. The fifth and last subtask consisted of the actual titration. The subject titrated the acid solution in the erlenmeyer flask, using the base contained in the buret. A color change, from clear to pink, signalled the neutralization of the acid within the flask. The subsequent volume of the base within the buret was recorded. This provided a measure of the accuracy by which all the solutions in the task were measured.

Any procedural errors during the full task were recorded. These errors included deviations from the task directions, equipment breakage, and chemical and water spills. Time stopped when the subject stopped the timer given to him to indicate completion of the task.

The titration task was performed twice more by the subject for a total of three trials. After the completion of the last trial, a questionnaire was given inquiring about the subject's perceptions regarding direction clarity, laboratory safety, apparatus use, stress, and task enjoyment (Appendix D). After completing the questionnaire, the experimenter reminded the subject to return next week for the second session and then told the subject not to talk to anyone about the experiment until all the data for the study had been collected.

Second session. Upon returning the following week, the subject was again read the general directions and precautions. The subject was told the procedures were similar to the last experiment and the risks involved were the same. The subject was then asked to re-familiarize himself with the equipment.

The same sequence used for the first session was used for the second session. Three trials were also performed, and the general session questionnaire given after the completion of the third trial. Additionally, a perceived risk questionnaire (referred to as the "final questionnaire") was given to the subject, covering the activities of both sessions (Appendix E).

Risk tendency. Wilde stated that "the role requirements within society appear to be such that we need some individuals to be risk takers... and others to be risk avoiders...." (Wilde, 1982b, p. 254). In order to explore the effects of risk tendency on RHT, the

TABLE 2

Breakdown of Titration Task

Subtask	Subtask Beginning	Subtask End
preparing solution "S"	initiation of task	returning solution "S" to table
adding acid "A"	returning solution "S" to table	finish adding acid "A" to solution "S"
adding water and indicator	finish adding acid "A" to solution "S"	returning indicator bottle to table
adding base "B"	returning indicator bottle to table	returning base "B" to table
performing titration	returning base "B" to table	writing buret vol. after titration task completion

Zuckerman's Sensation Seeking Scale, Version V (SSS; Appendix F) was administered at the completion of the second session.

The subject was then debriefed as to the true purpose of the study. The researcher fielded any questions the subject had at this time and answered those brought up previously. The subject was then paid, and again told not to discuss the experiment with anyone.

Experimental Design

The independent variable was the use of protective equipment. The four treatment condition levels created by the presence or absence of protective equipment during the two sessions were described above. The two control groups were those that had the protective equipment on both sessions or no equipment on both sessions. A repeated measures design was employed, with each subject performing three trials during each of two sessions. Session by Condition interactions were used to identify treatment effects. The three trials per session were used to study practice effects and intra-session differences. The experimental design in Figure 3 illustrates the four treatment condition levels, the two sessions and the three trials within each session.

The resulting data from the experiment were analyzed to determine whether the four hypotheses listed in the literature review were confirmed. Responses on the Session and Final Questionnaires were also analyzed to gauge perceptions of safety during the task. The questionnaire responses were correlated with each other and with the dependent measures.

Scores from Zuckerman's SSS were used to explore the effects of risk-taking tendencies on RHT. Multiple regression analyses were performed to determine the strength

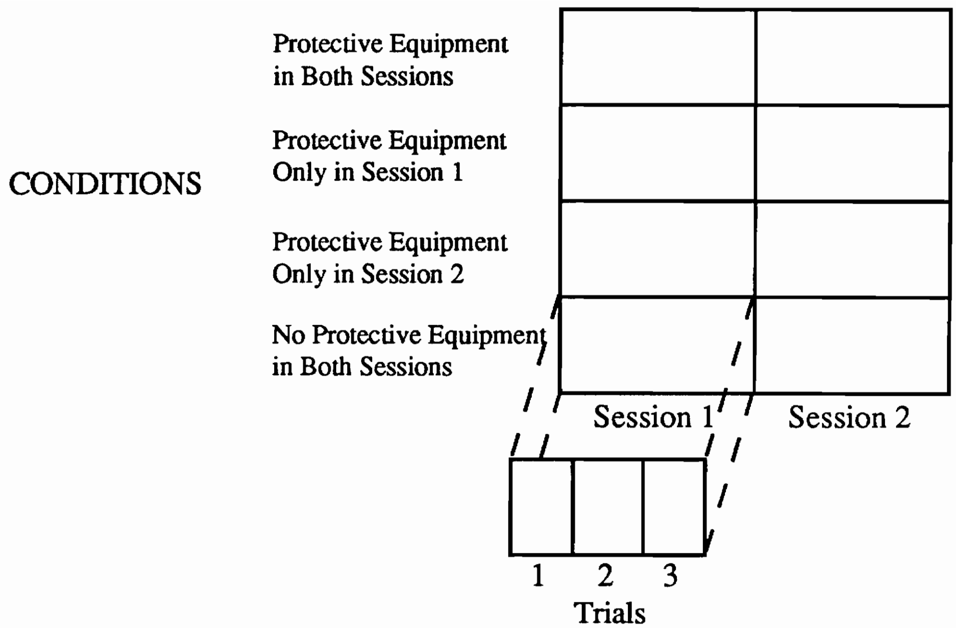

Figure 3: Experimental Design.

of risk tendency in predicting time and accuracy results. The analyses were used to assess the need for further studies into risk tendencies.

Dependent Variables. The dependent variables for this study are listed below:

1. Task completion times
2. Errors per trial
3. Error rate
4. Accuracy of the first and fifth subtasks

Total time for each trial was measured to the nearest second with a stopwatch. In addition, times were also kept for five separate sections per trial to determine where the largest differences in time occurred for the task. As previously mentioned, the errors per trial included deviations from the task directions, spills of solution "S", the acid, the base, and water, and apparatus drops. Error rate was computed by dividing the amount of errors committed per task with time required to complete the task. The performance accuracy of the first subtask was measured by taking the absolute difference between the obtained weight of solution "S" and the true weight of 5 ml of solution "S". The mean absolute differences for each group were then compared by session and by trial.

Initially, it was thought that accuracy could also be measured by determining the volume of base each subject used to neutralize the acid solution. The amount used by each subject would then be compared to the actual amount needed when the task is performed correctly. Unfortunately, it was determined that measurements of the base could not be done properly due to the dilute nature of the acid and base. Because the acid and base used were so weak, a sudden color change was not observed in the pilot study. Instead, a very gradual color change occurred. Although the experimenter provided the subjects with a "model neutralized solution," to which they could match the color of their solution, it was concluded that judging neutralization through color hues, as opposed to a sudden dramatic color change, would be too subjective to measure accuracy validly.

RESULTS

Elapsed Task Time per Trial

Figure 5 shows the mean elapsed time per trial for each experimental condition. The mean times for the four groups were all above nine minutes after the first trial. By the end of the third trial, the means had all fallen below nine minutes By the fifth trial, times had stabilized for all groups. The figure also shows that, on average, the groups with protective equipment in the first session performed more slowly than the groups without protective equipment through both sessions. The first three hypotheses, as they apply to the dependent variable of time, are as follows:

1) Subjects performing without protective equipment in the first session will perform in a less risky manner [at a slower pace] during the first session than subjects with protective equipment.
2) Subjects who switch from having protective equipment in the first session to not having protective equipment in the second session ("equipment - no equipment" condition) will perform in a less risky manner [at a slower pace] during the second session compared to those performing with protective equipment throughout both sessions ("equipment - equipment" condition).
3) Subjects who switch from not having protective equipment in the first session to having protective equipment in the second session ("no equipment - equipment" condition) will perform in a more risky manner [at a faster pace] during the second session compared to those performing without protective equipment throughout both sessions ("no equipment - no equipment" condition).

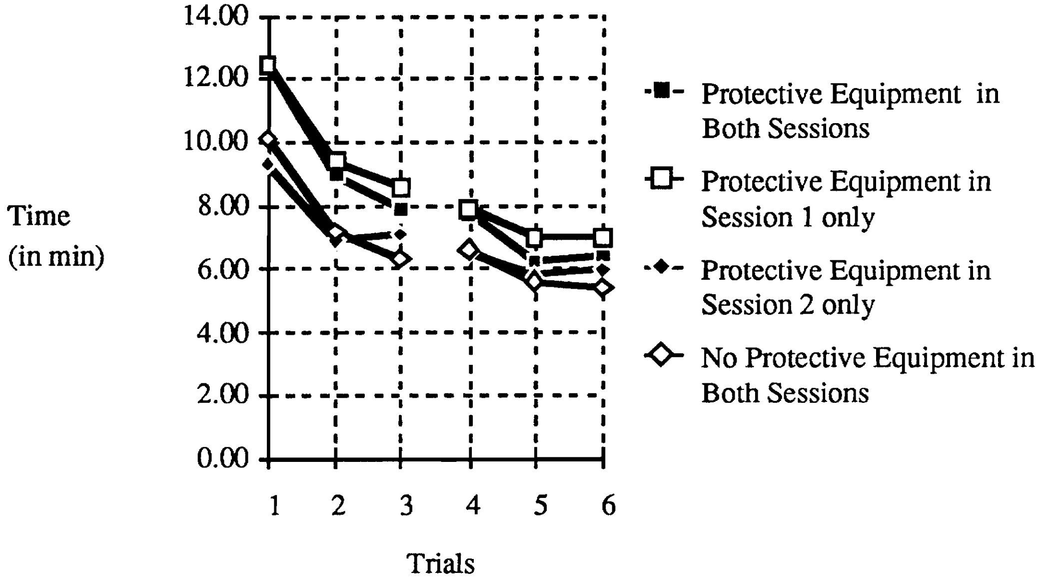

Figure 4: Mean completion times per trial.

To test Hypothesis 1, the treatment conditions were collapsed into those with protective equipment and those without protective equipment during the first session. A t-test was performed on the averaged trial data to test for a difference between the two groups during the first session. A separate variance t-test procedure was used when it was found that the variance of the two groups differed by a significant amount. The t-test equations are listed in Appendix J.

To test Hypotheses 2 and 3, second session data were subtracted from first session data to arrive at a difference score for each subject in the four conditions. A t-test on the difference scores of the group pairs specified in Hypotheses 2 and 3 were then performed. As with Hypothesis 1, a separate variance t-test procedure was used in situations where the variances between conditions differed significantly.

A 4x2 (4 Conditions x 2 Sessions) mixed factor analysis of variance (ANOVA) procedure was performed on the data to determine whether any of the four condition groups differed from the other in either or both session. The 4x2 ANOVA also determined whether a Condition x Session interaction occurred, indicating the expected separation in mean task completion times between the first and second session between the two experimental conditions.

A separate variance t-test was performed on the results of the first session to determine whether Hypothesis 1 was supported by the data. The results were significant in the direction opposite of that predicted by Hypothesis 1, with subjects who wore protective equipment performing at a slower pace than subjects who did not wear protective equipment during the first session, $t(23)=-2.62$, $p<.02$.

The t-test performed on the difference scores of those in the "equipment - equipment" and the "equipment - no equipment" conditions did not indicate a significant difference between the two with respect to task completion times in the second session. Therefore, Hypothesis 2 was not supported by the task time data.

Difference scores were also used to test Hypothesis 3. The t-test to determine whether a difference existed between subjects in the "no equipment - no equipment" condition and in the "no equipment - equipment" condition also did not indicate a significant difference.

The results of the 4x2 ANOVA indicated a significant session main effect, with elapsed task times generally decreasing from the first to the second session, $F(1,32)=91.12$, $p<.0001$ (Appendix G, Table 1). However, the condition main effect and the Condition x Session interaction were not significant.

Subtask Times

Subtask completion times were analyzed in a fashion identical to that of the full task. The purpose of taking subtask times was to determine whether behavioral changes occurred throughout the task or only at certain points. Graphical representations of the subtask times are given in Figures 5 - 9.

In general, the t-test results on first session subtask times were counter to the prediction made in Hypothesis 1. Subjects who wore protective equipment during the first session performed Subtask 3 at a significantly slower pace than subjects who did not wear protective equipment during the first session, $t(34)=-2.16$, $p<.04$. A separate variance t-test also revealed similar results for Subtask 5, $t(28)=-2.18$, $p<.04$. Although the results on subtask times for Subtasks 1, 2, and 4 were not significant, they too indicated that subjects who wore protective equipment performed at a somewhat slower pace than those who did not wear protective equipment (Subtask 1: $t(23)=-1.29$, $p<.21$; Subtask 2: $t(23)=$-

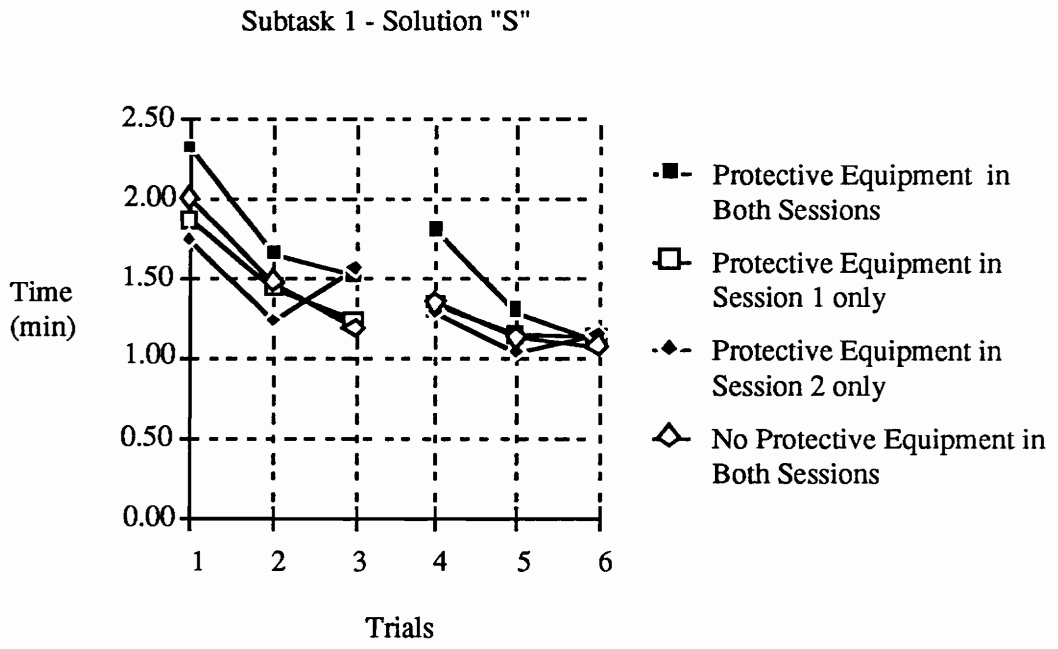

Figure 5: Mean time of completion per trial for Subtask 1.

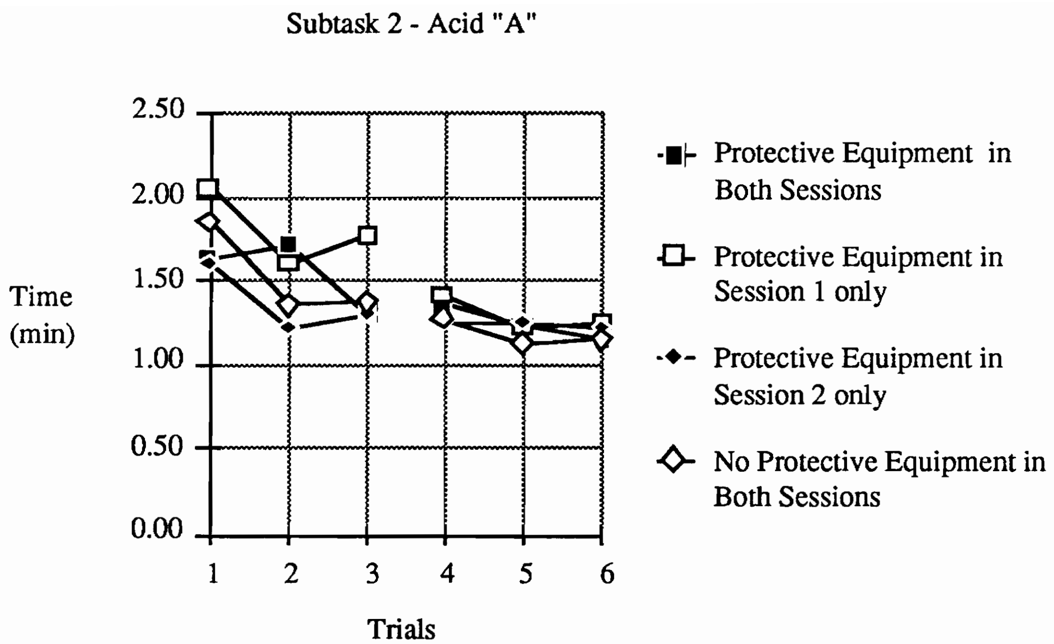

Figure 6: Mean time of completion per trial for Subtask 2.

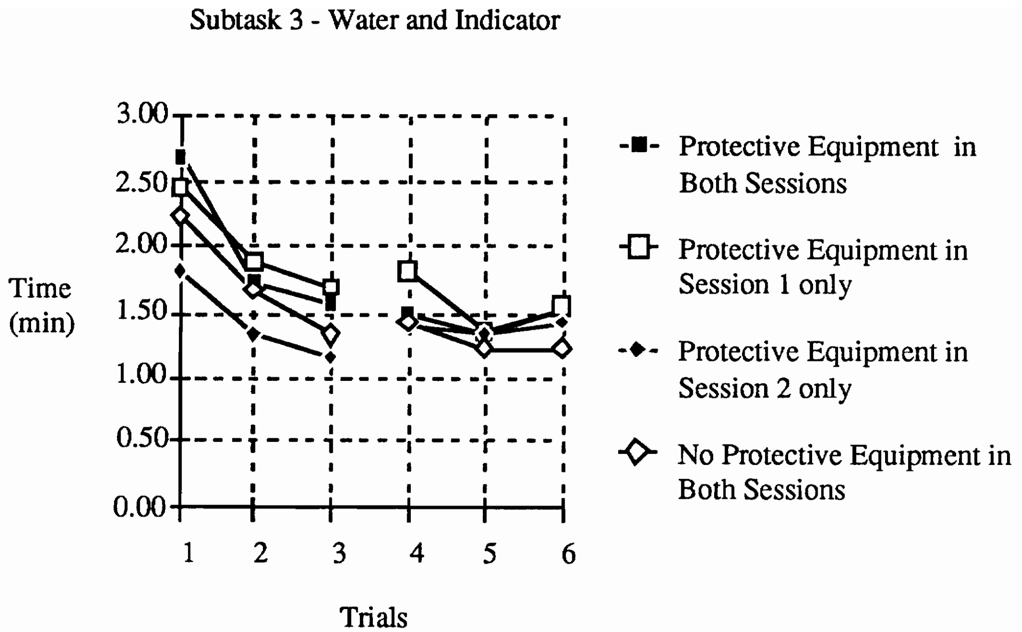

Figure 7: Mean time of completion per trial for Subtask 3.

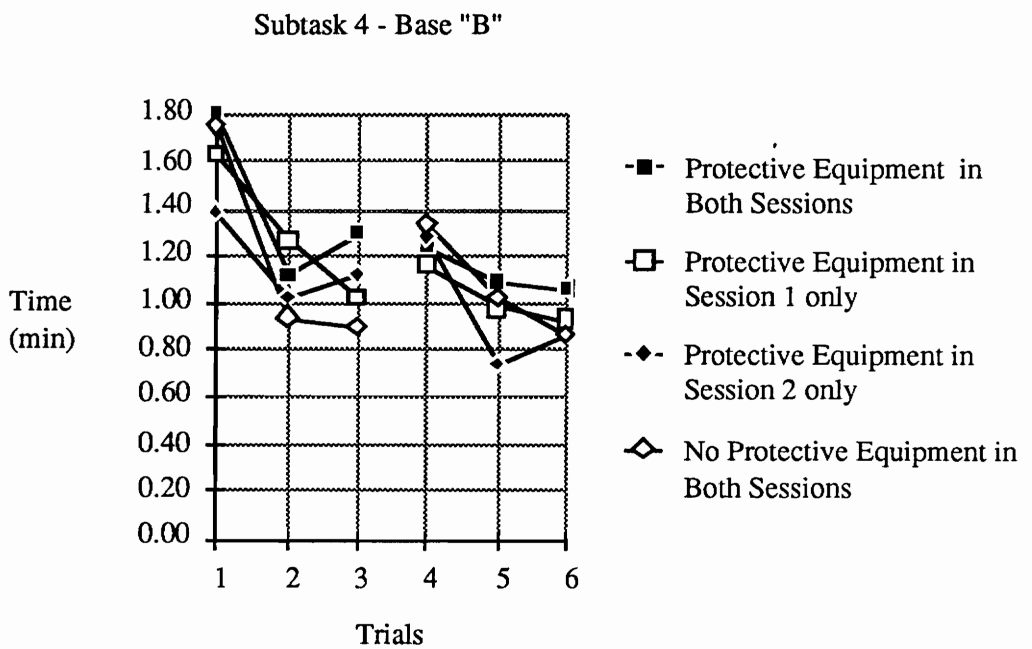

Figure 8: Mean time of completion per trial for Subtask 4.

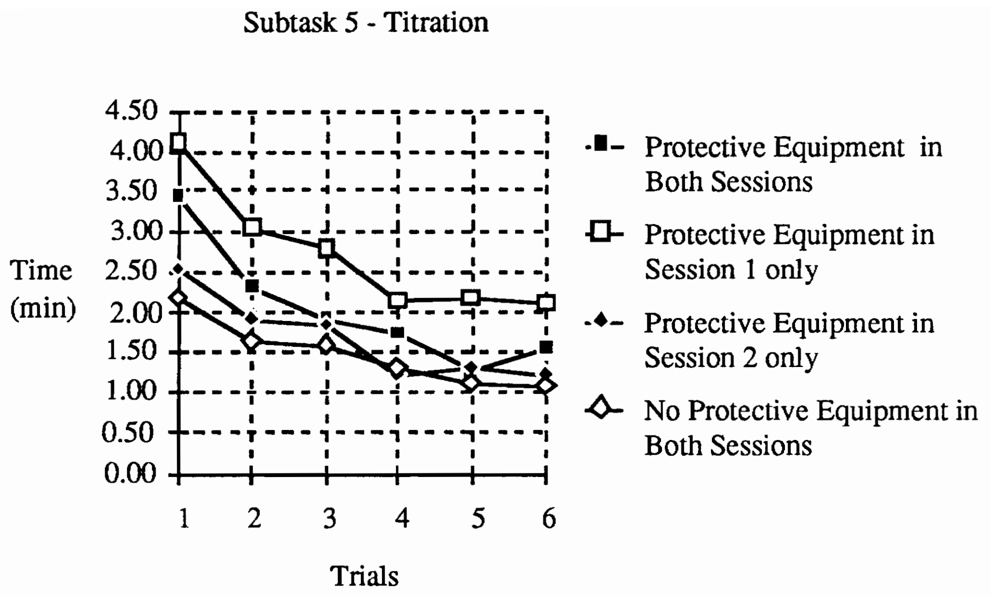

Figure 9: Mean time of completion per trial for Subtask 5.

1.36, p<.19; Subtask 4: t(34)=-1.56, p<.13). These results were all counter to Hypothesis 1.

Hypothesis 2 was not supported by the difference score data between those in the "equipment - equipment" condition and those in the "equipment - no equipment" condition. The t-test results for all five subtasks were not significant, indicating that the subtask times for both conditions decreased in a similar fashion from the first to the second session.

Hypothesis 3 was not supported by the difference score data between subjects in the "no equipment - equipment" condition and subjects in the "no equipment - no equipment" condition. The t-test results for all five subtask were not significant, but the results did indicate that the difference in pace of those in the "no equipment - equipment" condition was not as great from the first to the second session for Subtasks 2 and 3 compared to those in the "no equipment - no equipment" condition (Subtask 2: t(34)=-1.80, p<.09; Subtask 3: t(34)=-2.03, p<.06).

The 4x2 ANOVAs performed on the time data of the five subtasks all indicated session main effects, with task completion times consistently decreasing from the first to the second session (Appendix G, Tables 2-7). However, the 4x2 ANOVA for Subtask 5 also indicated a condition main effect, F(3,32)=3.41, p<.03. The Newman-Keuls Sequential Range Test revealed that, on average, subjects in the "equipment - no equipment" condition did not complete the task as quickly as the subjects in the "no equipment - no equipment" condition.

To understand better the significant condition main effect of the 4x2 ANOVA in Subtask 5, each session was analyzed separately using a 4x3 mixed factor ANOVA (4Conditions x 3 Trials; Appendix G, Tables 8-9). Analysis of the first session indicated that subjects in the "equipment - no equipment" condition took more time to complete the task than subjects in the "no equipment - no equipment" condition. However, the effect was not significant. A trial effect was found, with subtask completion times being significantly higher during the first trial than during the following two trials, $F(2,64)=12.22$, $p<.0001$. An analysis of the second session revealed a significant condition effect, $F(3,32)=4.36$, $p<.01$. A Newman-Keuls Sequential Range Test revealed that subjects in the "equipment - no equipment" condition took more time to complete the subtask than subjects in the other three conditions. This difference was in the direction predicted by Hypothesis 2.

Accuracy in Performing Subtask 1

The ability of subjects in each condition to obtain an accurate measure of the combined weight of solution "S" and the graduated cylinder was investigated using the identical analysis strategies previously mentioned. The first three hypotheses are restated below as they apply to this dependent variable:

1) Subjects performing without protective equipment in the first session will perform in a less risky manner [more accurately] during the first session than subjects with protective equipment.
2) Subjects who switch from having protective equipment in the first session to not having protective equipment in the second session ("equipment - no equipment" condition) will perform in a less risky manner [more accurately] during the second session compared to those performing with protective equipment throughout both sessions ("equipment - equipment" condition).
3) Subjects who switch from not having protective equipment in the first session to having protective equipment in the second session ("no equipment - equipment" condition) will perform in a more risky manner [less accurately] during the second session compared to those performing without protective

equipment throughout both sessions ("no equipment - no equipment" condition).

The t-tests performed on first session accuracy scores did not reveal any significant differences between those with and without protective equipment. The t-test on the difference scores for subjects in the "equipment - no equipment" and "equipment - equipment" conditions did not reveal a difference in accuracy. Although not significant, the separate variance t-test on the difference scores for subjects in the "no equipment - no equipment" and the "no equipment - equipment" conditions revealed that those in the "no equipment - equipment" condition improved their accuracy somewhat compared to those in the "no equipment - no equipment" condition, $\underline{t}(9)=-1.37$, $\underline{p}<.20$. The 4x2 ANOVA did not reveal any significant effects (Appendix G, Table 12).

Errors per Trial

Errors per trial, illustrated in Figure 11, was defined as the number of spills and procedural mistakes committed per trial. The hypotheses are restated as they apply to this dependent variable:

1) Subjects performing without protective equipment in the first session will perform in a less risky manner [commit less errors] during the first session than subjects with protective equipment.
2) Subjects who switch from having protective equipment in the first session to not having protective equipment in the second session ("equipment - no equipment" condition) will perform in a less risky manner [commit less errors] during the second session compared to those performing with protective equipment throughout both sessions ("equipment - equipment" condition).

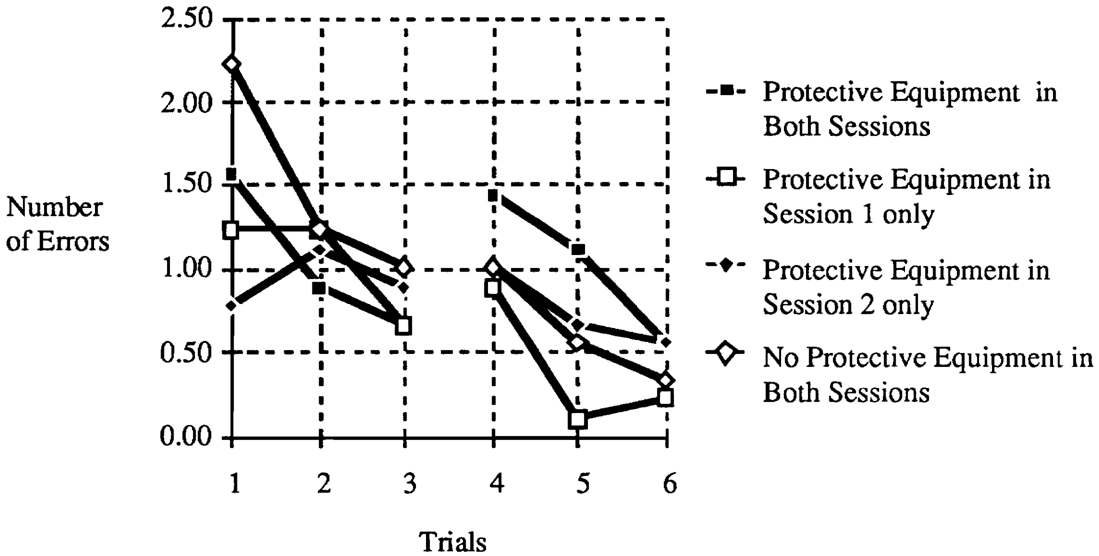

Figure 10: Mean number of errors per trial.

3) Subjects who switch from not having protective equipment in the first session to having protective equipment in the second session ("no equipment - equipment" condition) will perform in a more risky manner [commit more errors] during the second session compared to those performing without protective equipment throughout both sessions ("no equipment - no equipment" condition).

The t-test results on the first session data did not reveal a significant difference between those who performed with protective equipment and those who performed without protective equipment, lending no support to the prediction made in Hypothesis 1.

Due to the large differences in errors per trial during the first session (especially during Trial 1, among subjects who performed without protective equipment), the difference scores by which the t-tests for Hypotheses 2 and 3 were based were transformed into percent change scores using the formula:

$$\text{Percent Change} = \frac{(\text{Errors per Trial, Session 1}) - (\text{Errors per Trial, Session 2})}{\text{Errors per Trial, Session 1}}$$

The percentage change difference scores were checked for normality using the PROC UNIVARIATE procedure in SAS and were found to be normally distributed. A t-test on the difference scores of subjects in the "equipment - no equipment" and "equipment - equipment" conditions was not significant. Moreover, a t-test did not reveal a significant difference between the scores of subjects in the "no equipment - equipment" condition and those of subjects in the "no equipment - no equipment" condition.

A 4x2 ANOVA was performed on the error data. A significant Condition x Session interaction was found, $\underline{F}(3,32)=2.87$, $p<.05$, (Appendix G, Table 10). The Newman-

Keuls Sequential Range Test showed that the difference occurred during the second session, where subjects in the "equipment - no equipment" condition committed less errors, on average, than subjects in the "equipment - equipment" condition. This result was in the direction predicted by Hypothesis 2. A session main effect was also found, with errors decreasing from the first session to the second session, $\underline{F}(1,32)=13.97$, $p<.0007$.

Errors per Minute (Errors Rate)

Errors per trial were divided by their respective task times to determine the error rate for each trial (Figure 12). Hypothesis 4 stated that error rate will be equal across all conditions. It was expected that the subjects in all four conditions would have constant error rates throughout the two sessions since, according to the Risk Homeostasis Theory, the presence or absence of a safety measure will not affect the target risk level of an individual unless the measure is directed towards the *desire* to be safe.

Although not significant, the t-test performed on the first session error rates revealed that subjects who wore protective equipment committed less errors per minute than those who did not have protective equipment, $\underline{t}(34)=1.83$, $\underline{p}<.08$. Due to the large differences in first session error rates, the Percent Change formula was used again to determine the difference scores for error rate. The data were checked on SAS and found to be normally distributed. The t-test on the difference scores revealed that, though insignificant, subjects in the "equipment - no equipment" condition had a somewhat lower error rate than did subjects in the "equipment - equipment" condition, $\underline{t}(16)=1.30$, $\underline{p}<.21$. A significant difference did not arise from the t-test results between those in the "no equipment - no equipment" and "no equipment - equipment" conditions.

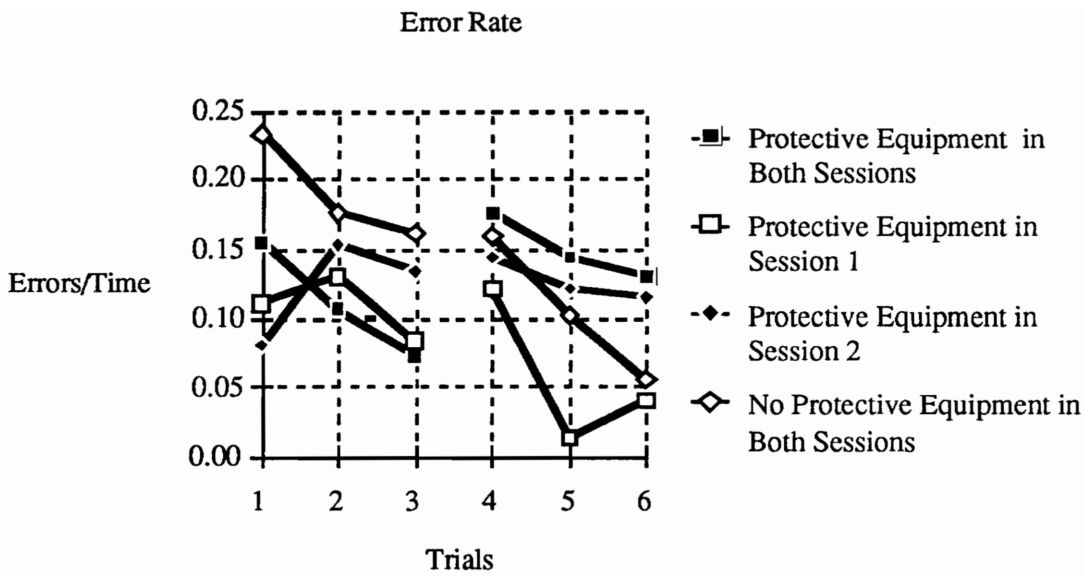

Figure 11: Errors per Minute (Error Rate) for each trial.

The 4x2 ANOVA indicated a significant condition effect, $F(3,32)=3.20$, $p<.04$ (Appendix G, Table 11). The Newman-Keuls Sequential Range Test showed that subjects in the "no equipment - no equipment" condition had a higher error rate than subjects in the "equipment - no equipment" condition. A marginal Condition x Session interaction was also present, $F(3,32)=2.77$, $p<.06$. The Newman-Keuls Sequential Range Test revealed that subjects in the "no equipment - no equipment" condition had a higher error rate than subjects in the other three conditions during the first session. Additionally, subjects in the "equipment - equipment" condition had a higher error rate than subjects in the "equipment - no equipment" condition during the second session. Both these results were contrary to the prediction given in Hypothesis 4.

Session Questionnaires

The first six items in the Session Questionnaire (Appendix D) were correlated with each other. A summary of significantly correlated items is given in Table 3. For the first session, four of the six items correlated positively with each other ($p.<.05$). Items dealing with the conciseness of the task directions (Item 1) and the adequacy of the protective equipment (Item 4) did not correlate with each other or with any other item. The significant correlations showed that subjects who believed that the apparatus were easy to use were likely not to feel as much anxiety or fatigue during the session, and were likely to believe the task could be completed safely and accurately within the recommended time of completion of nine minutes. Subjects who felt anxiety were likely to feel fatigued during the session.

Many correlations did not hold through the second session. Only two items were correlated with more than one other item. Anxiety felt during the task (Item 5) correlated positively with the adequateness of the safety precautions (Item 4; $r = .48$, $p = .003$) and

TABLE 3a

Correlation Analysis for Session 1 Questionnaire Items (Excluding Item 7)

Items	1	2	3	4	5	6
1	—	—	—	—	—	—
2	—	—	0.54**	—	0.33*	0.45**
3	—	0.54**	—	—	0.41**	—
4	—	—	—	—	—	—
5	—	0.33*	0.41**	—	—	0.49**
6	—	0.45**	—	—	0.49**	—

* $p < 0.05$
** $p < 0.01$

TABLE 3b

Correlation Analysis for Session 2 Questionnaire Items (Excluding Item 7)

Items	1	2	3	4	5	6
1	—	0.68**	—	—	—	—
2	0.68**	—	—	—	—	—
3	—	—	—	—	—	—
4	—	—	—	—	0.48**	—
5	—	—	—	0.48**	—	0.36*
6	—	—	—	—	0.36*	—

* $p < 0.05$
** $p < 0.01$

with the fatigue felt during the task (Item 6; r = .36, p = .03), while the ease at which the apparatus could be used (Item 2) correlated with the clearness and conciseness of the directions (Item 1; r = .68, p = .0001).

The treatment conditions for the first session were collapsed into those with protective equipment and those without protective. The data for the two groups were then independently correlated with the first session questionnaire responses. It was found that the task time data of those who had protective equipment during the first session were significantly correlated with how easily they felt the task could be completed safely and accurately in the recommended time (Item 3; r = -0.70, p=0.001), and the anxiety felt during the task (Item 5; r = -0.58, p=0.01). Trial errors were significantly correlated with the ease by which the apparatus could be used (Item 2; r=-0.51, p=0.03). Accuracy did not correlate with any of the questionnaire items. Task time data for those who did not have protective equipment during the first session were significantly correlated with Item 3 (r=-0.56, p=0.02) and with the feeling of no fatigue during the task (Item 6; r=-0.57, p=0.01). Trial error data correlated only with Item 3 (r=-0.50, p=0.04).

The treatment conditions for the second session were likewise collapsed into those with protective equipment and those without protective. Again, the data for the two groups were then independently correlated with the first session questionnaire responses. Task time data for those who did not have protective equipment during the first session were significantly correlated with Item 3 (r=-0.79, p=0.0001). There were no other significant correlations in the data of the collapsed groups.

Stepwise regressions were performed on each session questionnaire to predict task times, accuracy, errors and error rate for both sessions. Zuckerman's "Sensation Seeking Scale" (SSS) and "Thrill and Adventure Seeking" subscale (TAS) were also included as predictors. Only Item 3, whether the task could be completed safely and accurately in the recommended time, was found to predict task times significantly for the first session, R^2=.49, $F(1,34)=33.31$, $p<.0001$. Task times for the second session were also predicted significantly by Item 3, R^2=.48, $F(1,34)=31.41$, $p<.0001$.

The stepwise regression for Subtask 1 accuracy showed that only the TAS significantly predicted accuracy during the first session, R^2=.12, $F(1,34)=4.82$, $p<.04$. No combination of question items predicted significantly the accuracy for the second session. Trial errors in the first session were predicted significantly by the ease by which the apparatus could be used (Item 2), R^2=.17, $F(1,34)=6.96$, $p<.01$. None of the items from the second session questionnaire were able to predict significantly the errors in the second session. None of the items were able to predict the error rate for the first or second session.

In order to determine what effects the presence or absence of protective equipment had on the responses given on the Session 1 questionnaire, t-tests were performed on the means of the responses given by subjects who wore protective equipment and subjects who did not. None of the first session questionnaire items revealed a significant difference. However, the adequacy of the safety precautions (Item 4), showed a marginal effect, $t(34)=-1.98$, $p<.06$, with subjects performing without the protective equipment reporting the safety precautions as less adequate than subjects who wore the protective equipment.

Session 2 Questionnaire scores were subtracted from Session 1 Questionnaire scores to arrive at difference scores for each of the four conditions. The response differences between the first and second session questionnaires are illustrated in Figure 12. ANOVAs were conducted on the difference scores to determine if subject perceptions had changed from the first to the second session (Appendix H, Tables 1-6). A significant difference was found only for Item 4, $F(3,32)=9.46$, $p<.0001$ (Appendix H, Table 4). A Newman-Keuls Sequential Range Test indicated that the responses given by the subjects in

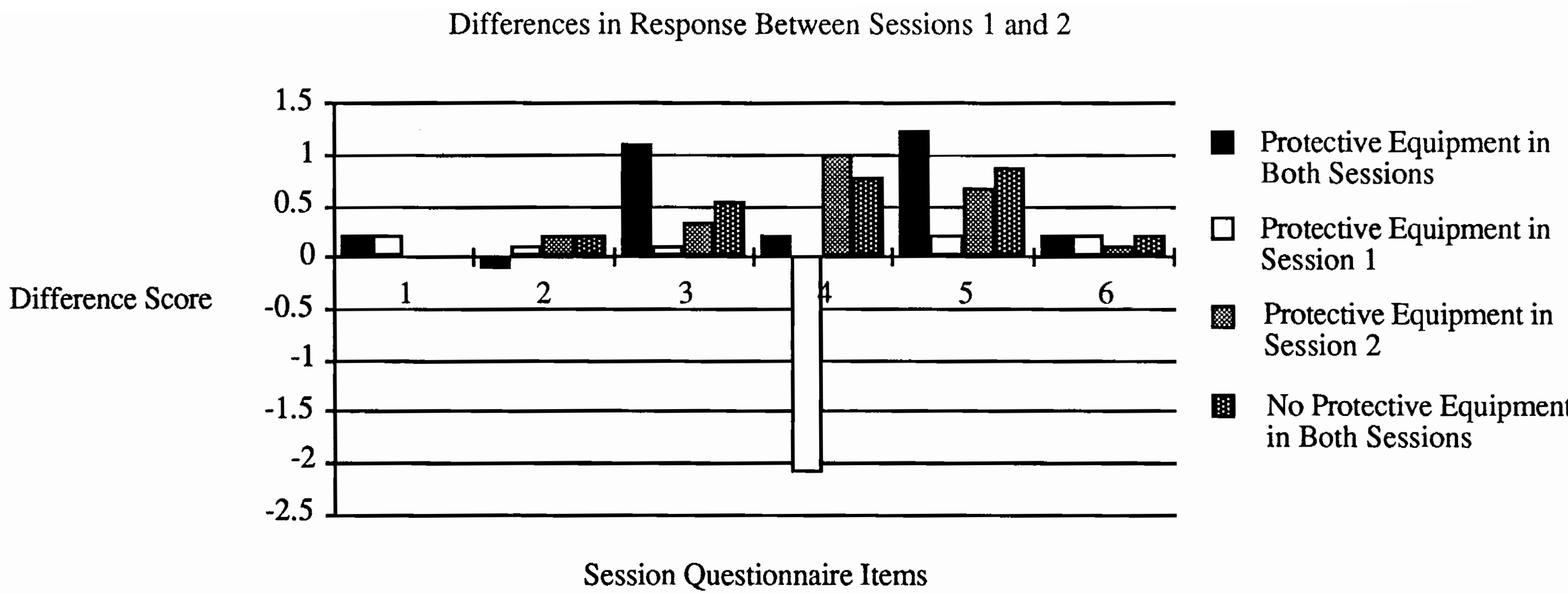

Figure 12: Differences in Response between the first and the second session questionnaire.

the "equipment - no equipment" condition were significantly different from the responses given by subjects in the other three conditions. Item 4 responses were an average of two points lower in the second session for subjects in the "equipment - no equipment" condition, while responses were higher in the second session for subjects in the other three conditions.

Item 7 of the Session Questionnaire was included only in the questionnaires of subjects wearing protective equipment in either or both of the two sessions. The purpose of the question item was to determine whether subjects felt that the protective equipment had hindered their performance during the session. In the first session, only subjects in the "equipment - equipment" and "equipment - no equipment" conditions were given the item. Surprisingly, the t-test performed on the two conditions revealed a difference, $t(16)=2.5$, $p<.02$, with those in the "equipment - equipment" condition feeling more hindered by the equipment than those in the "equipment - no equipment" condition (5.33 vs. 6.44, respectively).

In the second session, those in the "equipment - equipment" and "no equipment - equipment" conditions were given Item 7. The t-test did not reveal a difference between these two groups, $t(16)=-0.7$, $p<.49$, with those in the "equipment - equipment" condition having a mean of 5.78 and those in the "no equipment - equipment" condition having a mean of 5.44.

Final Questionnaire

The mean responses of each group for each item in the final questionnaire are illustrated in Figure 13. The items from the Final Questionnaire were correlated with each other. Table 4 summarizes the significantly correlated items. Changing one's pace when working without protective equipment (Item 1) did not correlate with any other item. This

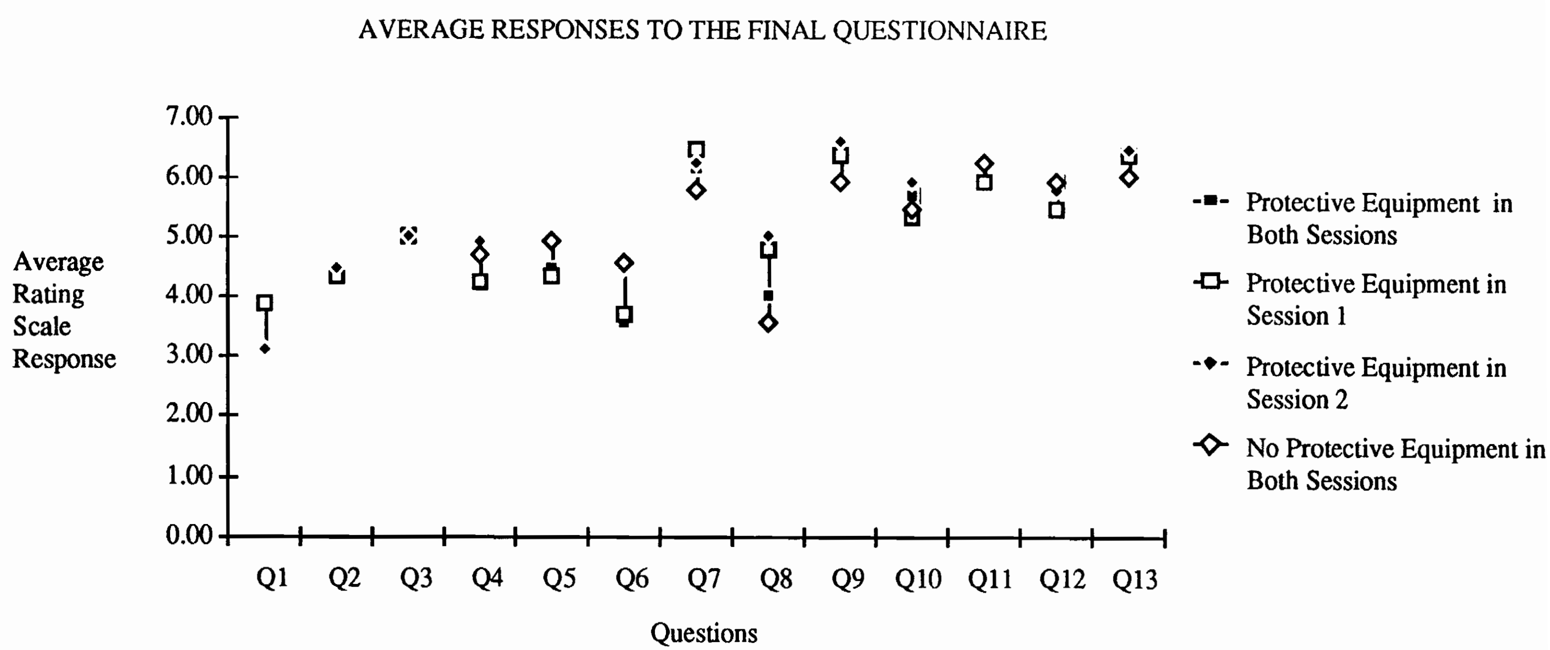

Figure 13: Average Responses of the Four Conditions to the Final Questionnaire.

TABLE 4
Correlation Analysis for Final Questionnaire Items

ITEM	1	2	3	4	5	6	7	8	9	10	11	12	13
1	—	—	—	—	—	—	—	—	—	—	—	—	—
2	—	—	-0.54**	0.42**	—	—	—	—	0.42**	0.47**	0.57**	0.62**	0.53**
3	—	-0.54**	—	-0.48**	—	-0.39*	—	—	-0.32**	—	—	-0.48**	-0.40**
4	—	0.42**	-0.48**	—	—	—	—	—	—	0.39*	0.40*	0.44**	0.47**
5	—	—	—	—	—	—	-0.34*	—	—	—	—	—	—
6	—	—	-0.39*	—	—	—	—	—	—	—	—	—	—
7	—	—	—	—	-0.34*	—	—	—	—	—	—	—	—
8	—	—	—	—	—	—	—	—	—	—	—	—	—
9	—	0.42**	-0.32*	—	—	—	—	—	—	0.53**	0.61**	0.47**	0.47**
10	—	0.47**	—	0.39**	—	—	—	—	0.52**	—	0.72**	0.77**	0.67**
11	—	0.57**	—	0.4*	—	—	—	—	0.61**	0.72**	—	0.70**	0.61**
12	—	0.62**	-0.48**	0.44**	—	—	—	—	0.47**	0.77**	0.70**	—	0.78**
13	—	0.53**	-0.4**	0.47**	—	—	—	—	0.47**	0.67**	0.61**	0.78**	—

* $p < 0.05$
** $p < 0.01$

was also the case with Item 8, the perceived benefits with the pace chosen. Perceptions of safety when performing without protective equipment (Item 2) correlated negatively with perceptions of safety when performing with protective equipment (Item 3, $r = -.54$, $p = .0007$); positively with the perceptions of safety associated with the chemicals used in the experiment (Item 4, $r = .42$, $p = .01$); positively with subjects' overall judgment of their pace (Item 6, $r = .31$, $p = .06$); and positively with perceptions of safety associated with each subtask (Items 9 - 13). Perceptions of safety when performing the task with protective equipment (Item 3) correlated negatively with Item 4 ($r = -.48$, $p = .003$); negatively with Item 6 ($r = -.39$, $p = .02$); negatively with perceptions of safety when weighing the powdered acid (Item 9, $r = -.31$, $p = .06$); negatively with the safety felt when preparing the base for titration, Item 12 ($r = -.48$, $p = .003$); and negatively with the safety felt while performing the actual titration, Item 13 ($r = -.40$, $p = .016$). Item 4 was correlated positively with the perception of safety associated with Subtasks 2 through 5. Changes in the perception of safety due to task repetition (Item 5) correlated negatively with the concern given to accuracy while performing the task (Item 7, $r = -.33$, $p = .05$). All items dealing with perceptions of safety associated with each subtask (Items 9 - 13) correlated positively with each other.

The 13 items of the Final Questionnaire, as well as the TAS and SSS, were then used as predictors of task times for the first and second sessions. Task times for the first session were significantly predicted only by the subjects' judgment of their pace (Item 6), $R^2=0.31$, $\underline{F}(1,34)=15.19$, $p<.0004$. The task times for the second session were, again, significantly predicted by Item 6, $R^2=0.31$, $\underline{F}(1,34)=10.78$, $p<.002$.

Accuracy for the first session was significantly predicted by the subjects' concern for accuracy while performing the task (Item 7), $R^2=0.17$, $\underline{F}(1,34)=6.98$, $p<.01$. Accuracy for the second session was not significantly predicted by any final question item.

Similarly, errors for both the first and second sessions were not significantly predicted by any final question item. Error rate for the first session was significantly predicted by Item 6, R^2=0.31, $F(1,34)$=8.12, $p<.007$. Error rate for the second session was not significantly predicted by any final question item.

A 2x2 (Session 1, equipment / no equipment x Session 2, equipment / no equipment) ANOVA was conducted on Item 6 to determine whether subjects in the four conditions differed in how fast they believed they worked during the study (Appendix I). Although the differences were not significant for each of the four conditions, subjects who wore protective equipment during the first session reported going at a slower pace than did subjects who did not wear protective equipment during the first session, $F(1,32)$=5.60, $p<.02$.

Risk Tendency as a Predictor of Pace and Error

As mentioned in previously, subjects' risk and sensation seeking tendencies, as measured by Zuckerman's Sensation Seeking Scale (SSS; Zuckerman, 1979) were not a reliable predictor of task times, accuracy or error, overall. The Thrill and Adventure Seeking Subscale (TAS) did significantly predict first session accuracy scores, R^2=.12, $F(1,34)$=4.82, $p<.04$, but the strength of prediction was of little practical importance.

DISCUSSION

The purpose of this experiment was to determine whether aspects of the Risk Homeostasis Theory could be detected in a laboratory environment. The aspects of the theory investigated were the existence of risk compensation behavior and of a homeostatic risk regulating mechanism. In general, the results indicated that subjects in the experimental conditions did perceive changes in the risk associated with the titration task between the first and second session. Yet, risk compensation behavior, as stated in the Risk Homeostasis Theory, was not found in the analyses of the dependent variables. A constant error rate, indicating a homeostatic regulating mechanism, was also absent.

Summary of Findings for Hypotheses 1 - 3

Results on task completion times, accuracy, and task errors failed to support the between-subject hypothesis that those without protective equipment would perform in a less risky manner than those with protective equipment during the first session. There were no significant differences in the amount of errors per trial, nor in the performance accuracy of Subtask 1. There were significant effects detected in the task time data and in the time data of Subtasks 3 and 5. However, the results were contrary to what would have been expected from Wilde's theory, with those who wore protective equipment having performed the task and subtasks at a significantly slower pace than those who did not wear the equipment.

Although the data did not reveal a compensation effect in the first session, responses to the first session questionnaire indicated that subjects who did not wear protective equipment in the first session perceived somewhat more risk than subjects who wore protective equipment. Additionally, two subjects who performed the first session without protective equipment expressed their desire for gloves in the comment section of the questionnaire.

Hypothesis 2 was not sufficiently supported by the data. The 4x2 ANOVAs for trial errors and for Subtask 5 indicated that subjects in the "equipment - no equipment" condition had compensated in the predicted directions, relative to the performance of the other three groups. However, the t-scores indicated that subjects in the "equipment - no equipment" condition did not perform the task or any of the subtasks at a significantly slower pace during the second session compared to subjects in the "equipment - equipment" condition. A large difference in trial error mean scores did exist between the first and second session, but the large variance in the difference scores for both groups washed out a significant effect. Accuracy was consistent throughout the sessions for both groups. However, the absence of protective equipment during the second session did cause subjects in the "equipment - no equipment" condition to feel much less safe, as indicated by the highly significant effect found in the analysis of the Item 4 data in the session questionnaire.

Hypothesis 3 was also not supported by the t-tests on the difference score data. All subjects in the "no equipment - equipment" condition responded on Item 4 of the second session questionnaire that they felt "Very Safe." However, these subjects were not found to perform the subtasks at a significantly faster pace, perform Subtask 1 less accurately, or commit significantly more errors per trial compared to subjects in the "no equipment - no equipment" condition. In fact, marginal t-scores on accuracy and subtask completion times tended to show effects contrary to Hypothesis 3, with subjects in the "no equipment - equipment" condition increasing in accuracy and posting a smaller difference in task completion times from the first to the second session than those in the "no equipment - no equipment" condition.

Interpretation of Findings for Hypotheses 1 - 3

Several explanations can be given for the lack of support for the risk compensation aspect of the Risk Homeostasis Theory. One involves the seeming complexities of predicting subject reaction towards changes in perceived risk. During the debriefing, one subject from the "equipment - equipment" condition commented that wearing the protective equipment during the task made him more aware of potential dangers, thereby causing him to act cautiously during the first session. The negative correlation between the lack of anxiety (Item 5 in the session questionnaires) and task completion times for subjects who wore protective equipment during the first session further suggest that feedback from the protective equipment may not have consistently shifted behavior in the direction predicted by Wilde's theory. This would explain the presence of somewhat large variances in time data.

It may also be possible that subjects in both experimental conditions perceived a change in risk but did not feel the change large enough to warrant a behavioral adjustment. As indicated by the mean responses to Item 4 of the final questionnaire, subjects perceived the chemicals as somewhat safe. Mean responses to Item 5 of the final questionnaire indicated that repeating the task generally caused subjects to judge the chemicals as much safer. Moreover, subjects reported feeling generally safe while performing the five subtasks (Items 9-13 of the final questionnaire).

The explanation above also does not support the Risk Homeostasis Theory. Wilde (1985) mentioned that it was possible for a protective measure below the threshold level not to produce a compensatory response, but those above the threshold level will be noticed and will produce a compensatory action. It was obvious from the session questionnaire of the present study that subjects in the experimental conditions did perceive a change in risk, yet they did not compensate for this shift in risk.

Another possibility is that the protective equipment may have confounded the results of the study. In the comment section of the first session questionnaire, two subjects from the "equipment - equipment" condition reported that the goggles had a tendency to "fog up" during the task. The problem with goggles fogging was not discovered during the pilot study. This may be due to the somewhat warmer outside temperatures during mid-November when the pilot study was performed (in contrast, the actual study took place from mid-January to early February). Additionally, one subject mentioned that the gloves made picking up the stirring rods somewhat difficult.

Although there is the possibility that the performance of some subjects may have been affected by the use of the protective equipment, it is not clear whether the equipment masked potentially significant results or caused unexpected ones. A case in point is that while problematic equipment may explain why subjects in the "equipment - no equipment" condition performed subtask 5 at a significantly slower pace than subjects in the "no equipment - no equipment" condition during the first session, it does not explain why subjects in the "equipment - no equipment" condition continued to perform subtask 5 at a significantly slower pace than subjects in the remaining three conditions during the second session. Equally important, the mean responses to Item 7 of the first and second session questionnaires suggested that the protective equipment did not hamper performance; and the average response to Item 1 of the final questionnaire indicated that subjects in the experimental groups believed that working without the protective equipment caused them to decrease their pace somewhat.

Summary and Interpretation of Findings for Hypothesis 4

Wilde predicted that accident rates per unit time would remain at a constant level unless affected by a motivational safety measure. Based on the assumptions given in the Literature Review concerning task severity, it was predicted that error rate would be equal across all four conditions. The results did not support this expectation.

However, upon further review, the error rate results should not have come as a surprise in the short-term situation where accident severity is assumed to be zero. For example, if the error rate for subjects in the "equipment - no equipment" condition were to stay at the same level from the first to the second session, one of three things must occur. First, subjects must not exhibit behavioral adjustments through both task time and errors. Second, if compensation was exhibited through task time (i.e., task time increases), reverse compensation must occur through the number of errors made (more errors committed). Third, if compensation was exhibited through errors (error decrease), reverse compensation must occur through task time (shorter completion times). Obviously, the occurrence risk compensation behavior, as predicted in Hypotheses 1 through 3, could not allow a constant error rate to occur in a zero severity situation.

This does not mean that Hypothesis 4 was incorrectly stated. However, it does point to the importance of severity in the Risk Homeostasis Theory equation for accident rate. Since the objective risk of the present experiment was negligible, accident severity for all types of errors was assumed to be identical and was dropped from the accident rate equation. In effect, this meant that whether or not a subject was wearing protective equipment, the consequence of the error was the same. Even though the lack of objective risk made this assumption acceptable, it did not take into account subject perceptions of error severity.

In hindsight, trial errors committed during the pilot study should have been categorized. Each category should have been given a perceived severity rating, taking into account the presence or absence of protective equipment. This perceived severity rating, with which subjects in the actual study would be familiarized, should have then been used to weigh the errors committed per minute to arrive at a perceived rate of errors.

Most, if not all researchers studying risk homeostasis behavior have made the mistake of attempting to explain the homeostatic mechanism specified in Wilde's theory by taking into account only one level of severity. Researchers have usually compared fatalities before and after the implementation of a safety measure to argue for or against the presence of risk homeostasis. This was done because data on fatalities were more reliable than data covering all other levels of accident severity. However, any knowledge of critical incidence ratios makes it clear that accidents resulting in a fatality represent only a small portion of total accidents, the majority of which are non-fatal. Only by taking into account all degrees of severity can a homeostatic mechanism be demonstrated or not supported.

Session and Final Questionnaires

Correlations on the first session questionnaire items did not reveal the expected relationships between the perceived adequateness of the safety precautions (Item 4) and other task perceptions during the first session. However, the correlations did seem to show a relationship between the ease by which the chemistry apparatus could be used

(Item 2) and the lack of fatigue (Item 6) and anxiety felt (Item 5). It is hard to determine whether difficulties with the chemistry apparatus caused the feelings of anxiety and fatigue, or whether feelings of fatigue and anxiety led to difficulties with the apparatus, but the comment section of the session questionnaire suggested that one possible cause of the anxiety may have been the presence of the video camera more than the titration equipment. Four subjects stated that they felt uncomfortable being video-taped during the task. The ability of the task to be completed within the recommended time limit of nine minutes (Item 3) was also correlated with Item 5, suggesting that the inability to meet the time limit may have also contributed to feelings of anxiety.

Significant correlations in the second session questionnaire indicated that subjects who found the directions clear and concise (Item 1) also found the apparatus easy to use (Item 2). Item 5 also correlated with the adequateness of the safety precautions (Item 4), suggesting that the anxiety may have been caused by concerns with safety. No comments were made regarding the video camera during the second session.

Item 3 in both the first and second session questionnaires was the only item to predict task times significantly during both sessions. These regression equations suggest that the degree of concern given to safety and accuracy may been the main determinant of task times. These regression equations also confirm that risk was perceived by the subjects of this study. While Subtask 1 accuracy for the first session was significantly predicted by the TAS, and errors for the second session was significantly predicted by Item 2, the prediction power of both were too low to be of practical significance.

Responses to Item 6 of the final questionnaire showed that subjects were quite accurate in judging their pace. On average, those who wore protective equipment during the first session judged their pace as being somewhat slow, while those who did not wear protective equipment during the first session judged their pace as being somewhat fast. Figure 5 in the Results section shows that those with protective equipment during the first session did perform significantly slower in both sessions, on average, than those who did not have protective equipment during the first session.

It can be inferred from the negative correlation between Items 5 and 7 of the final questionnaire that those who felt that repeating the task made the chemicals seem safer (Item 5) also were less concerned with accuracy as they proceeded through the task iterations. The correlation between Item 2 and the other questionnaire items suggested that those who perceived the subtasks as safe (Items 9-13) and the solutions as not dangerous (Item 4) also felt relatively safe performing the task without protective equipment.

The Final Questionnaire regression equations for task times indicated that the subjects' judgement of their pace (Item 6) was the best predictor of task completion times. This suggested that the subjects were conscious of their pace as they performed the task. Concern with accuracy while performing the task (Item 7) was the best predictor of first session Subtask 1 accuracy scores, indicating that some subjects were also consciously aware of accuracy during their performance. However, the low strength of prediction made this finding of little practical importance.

Comparison with Similar Studies

The results of this study can be put into perspective by comparison with other studies on risk compensation and risk homeostasis. The absence of an effect during the first session of the presence of study was consistent with the findings of O'Neil, Lund, and Zador (1985), who did not find evidence of driver risk compensation behavior between subjects after seat belt laws were instituted in England and Canada. The researchers twice detected decreases in travel speed, as opposed to only one increase in speed, following the implementation of the seat belt laws of both countries. Furthermore, they found no change in following headways, turning headways, or responses to yellow signals. Similarly, the first session results of the present study were either inconclusive or contradictory to the Risk Homeostasis Theory.

On the other hand, the absence of risk compensation behavior in the within-subject analysis was not consistent with the results reported by Streff and Geller (1988), who did measure a significant compensation effect. Streff and Geller (1988) concluded from their results that risk compensation occurred only when a subject was able to contrast an unsafe situation with a safer situation. In contrast, while the expected perceived change in risk was detected in the present study, compensation was not measured in any of the three dependent variables.

Conclusion

The results of this study showed that a shift in risk through the presence or absence of non-motivational protective measures may cause changes in behavior not necessarily in the directions expected from the Risk Homeostasis Theory, suggesting that human behavior in the face of risk may be more complex than Wilde and other compensation theorists believe.

The present study failed to support the existence of a homeostatic risk regulating mechanism. However, the technique by which the existence of the mechanism was measured must be improved before conclusive statements can be made.

This study did demonstrate that perceived risk can be manipulated and measured while minimizing objective risk, if not eliminating it altogether. This has been an obstacle in the testing of Wilde's theory. Although a study such as this may not be possible under all settings, it is promising in the research where the perceived risk can temporarily mask the true objective risk. A staged "accident", designed to exemplify the "consequences" of an error, may add to perceptions of risk.

Another option is to devise a study in which the objective risk is acceptable but not desirable. In this way, a stable change in behavior can be observed and "losses per unit time" analyzed to determine whether a constant level of risk is maintained. Using the present study as an example, the experimenter could warn subjects they must finish the task as fast as possible and in a safe manner, or be subject to a form of undesirable punishment, such as consuming varied amounts of a bad tasting but safe liquid, depending upon the severity of the error. Such an experiment has an advantage over monetary reward or loss because an error results in an acceptable form of physical discomfort, which may be more easily equated to a true safety-related situation. The experimenter may also find ways to increase the realism of the situation by asking the subjects to imagine each spoon of the bad tasting liquid as a more damaging injury.

The unfortunate possibility exists that the results of the present study were confounded by fogging goggles. Additionally, the the informed consent form may have given subjects the impression that the task was relatively safe. Finally, due to University regulations concerning the use of acids by students, site selection was limited to rooms not completely adequate for the study (e.g., the absence of one-way mirrors needed to conceal the video equipment).

Future Studies

Besides investigating the existence of risk compensation and of a homeostatic risk regulating mechanism, another aspect of Wilde's theory which *may* be possible to study concerns whether motivational measures safety are actually better than nonmotivational safety measures in reducing the accident rate, as defined by Wilde. For example, it could be investigated whether an increased probability of penalty for bad driving (a motivational measure that reduces the benefits of risky behavior) reduces the accident rate to a greater extent than the use of safety belts (nonmotivational measure). However, before a study such as this can take place, it must be determined that these two measures are similar in their ability to reduce the accident rate.

Another topic for research may be the notion of risk transfer from one mode of activity to another. Such a study would require subjects to perform two separate tasks. The experimenter would raise the perceived risk of one task and would then attempt to determine whether the subjects transfer risky behavior from one task to another. Such an experiment would be a major undertaking, since risk taking behavior would have to be equated from one task to the other. Other subtheories, such as the ability of risk to transfer from one subpopulation to another may well be impossible to test.

The present study has demonstrated that it is possible to test aspects of the Risk Homeostasis Theory, such as behavior compensation and a homeostatic risk regulating mechanism for *individuals*. However, echoing the comments of McKenna and Evans, it is up to Wilde to show that this theory is worthwhile to be researched further. This may mean simplifying, altering, or altogether eliminating certain subtheories. What matters most is the advancement of ideas, and much would be learned whether or not the Risk Homeostasis Theory is eventually supported or refuted.

BIBLIOGRAPHY

Adams, J. G. (1988). Risk homeostasis and the purpose of safety regulation. Ergonomics, 31(4), 407-428.

Cole, G. A., & Withey, S. B. (1982). The risk of aggregation. Risk Analysis, 2(4), 243-248.

Evans, L. (1985). Human behavior feedback and traffic safety. Human Factors, 27(5), 555-576.

Evans, L. (1986). Comments on Wilde's notes on "risk homeostasis theory and traffic accident data". Risk Analysis, 1(6), 103-107.

Evans, L. (1986). Risk homeostasis theory and traffic accident data. Risk Analysis, 1(6), 81-94.

Finn, P., & Bragg, B.W. (1986). Perception of the risk of an accident by young and older drivers. Accident Analysis and Prevention, 18(4), 289-298.

Fuller, R. (1984). A conceptualization of driving behaviour as threat avoidance. Ergonomics, 27(11), 1139-1155.

Graham, J. D. (1982). On Wilde's theory of risk homeostasis. Risk Analysis, 2(4), 235-242.

Herms, B. F. (1972). Pedestrian crosswalk study: accidents in painted and unpainted crosswalks. Highway Research Record, 406, 1-13.

Howarth, C. I. (1988). The relationship between objective risk, subjective risk and behaviour. Ergonomics, 31(4), 527-535.

Janssen, W., & Tenkink, E. (1988). Risk homeostasis theory and its critics: time for an agreement. Ergonomics, 31(4), 429-433.

Janssen, W. H., & Tenkink, E. (1988). considerations on speed selection and risk homeostasis in driving. Accident Analysis And Prevention, 20(2), 137-142.

Jocksch, H. C., (1976). Critique of Sam Peltzman's study: the effects of automobile safety regulation. Accident Analysis and Prevention, 8, 129-137.

McKenna, F. P. (1982). The human factor in driving accidents: an overview of approaches and problems. Ergonomics, 25(10), 867-877.

McKenna, F. P. (1985). Do safety measures really work? An examination of risk homoeostasis theory. Ergonomics, 28(2), 489-498.

McKenna, F. P. (1985). Evidence and assumptions relevant to risk homoeostasis. Ergonomics, 28(11), 1539-1541.

McKenna, F. P. (1988). What role should the concept of risk play in theories of accident involvement? Ergonomics, 31(4), 469-484.

Molen, H. H. Van Der, & Botticher, A. M. (1988). A hierarchical risk model for traffic participants. Ergonomics, 31(4), 537-555.

Naatanen, R., & Summala, H. (1976). Road-User Behavior and Traffic Accidents. Amsterdam and New York: North-Holland/American Elsevier.

O'Neil, B., Lund, A.K., & Zador, P. (1985). Mandatory belt use and driver risk taking: an empirical evaluation of the risk-compensation hypothesis. Human Behavior and Traffic Safety, 93-107.

Oppe, S. (1988). The concept of risk: a decision theoretic approach. Ergonomics, 31(4), 435-440.

Partyka, S. C. (1984). Simple models of fatality trends using employment and population data. Accident Analysis and Prevention, 16(3), 211-222.

Peltzman, S (1975). The effects of automobile safety regulation. Journal of Political Economy, 85(4), 677-725.

Robertson, L. S. (1977). A critical analysis of Peltzman's 'The effects of automobile safety regulation'. Journal of Economic Issues, 11, 587-600.

Shannon, H. S. (1986). Road-accident data: interpreting the British experience with particular reference to the risk homoeostasis theory. Ergonomics, 29(8), 1005-1015.

Slovic, P., & Fischhoff, B. (1982). Targeting risk. Risk Analysis, 2(4), 227-234.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1978). Accident probabilities and seat belt usage: a psychological perspective. Accident Analysis and Prevention, 10, 281-285.

Streff, F. P., & Geller, S. E. (1988). An experimetal test of risk compensation: between-subject versus within-subject analyses. Accident Analysis and Prevention, 20(4), 277-287.

Summala, H. (1988). Risk control is not risk adjustment: the zero-risk theory of driver behaviour and its implications. Ergonomics, 31(4), 491-506.

Taylor, D. H. (1964). Driver's galvanic skin response and the risk of accident, Ergonomics, 7, 439-451.

Van der Molen, H. H., & Botticher, A. M. T. (1988). A hierarchical risk model for traffic participants. Ergonomics, 31(4), 537-555.

Wilde, G. J. (1982). Critical issues in risk homeostasis theory. Risk Analysis, 2(4), 249-258.

Wilde, G. J. (1982). The theory of risk homeostasis: implications for safety and Health. Risk Analysis, 2(4), 209-225.

Wilde, G. J. (1984). Evidence refuting the theory of risk homoeostasis? A rejoinder to Frank P. McKenna. Ergonomics, 27(3), 297-304.

Wilde, G. J. (1985). Assumptions necessary and unnecessary to risk homoeostasis. Ergonomics, 28(11), 1531-1538.

Wilde, G. J. (1986). beyond the concept of risk homeostasis: suggestions for research and application towards the prevention of accidents and lifestyles-related disease. Accident Analysis And Prevention, 18(5), 377-401.

Wilde, G. J. (1986). Notes on the interpretation of traffic accident data and of risk homeostasis theory: a reply to L. Evans. Risk Analysis, 1(6), 95-102.

Wilde, G. J. (1988). Risk homeostasis theory and traffic accidents: propositions, deductions and discussion of dissension in recent reactions. Ergonomics, 31(4), 441-468.

Wilde, G. J. (1991). Economics and accidents: a commentary. Journal of Applied Behavior Analysis, 24(1), 81-84.

Wilde, G. J., Claxton-Oldfield, S. P., & Platenius, P. H. (1985). Risk Homeostasis in an experimental context. Human Behavior and Traffic Safety. 119-142.

Wilde, G. J., & Murdoch, P. A. (1982). Incentive systems for accident-free and violation-free driving in the general population. Ergonomics, 25(10), 879-890.

APPENDIX A

Informed Consent Form

INFORMED CONSENT FORM

1. The purpose of this study is to investigate safety tendencies in a laboratory setting.

   You will be asked to perform a series of six acid-base titrations. Three titrations will be done today and the remaining three will be completed on the same day next week. Be sure of your availability on both days. Inform the researcher if you foresee any problems.

2. You will be paid at a flat rate of $15.00 an hour. The estimated total time needed to finish both sessions is 2.5 hours. Although time is also of concern, the ability to perform accurately and safely is of primary importance.

3. There are some risk involved to which you expose yourself in volunteering for this research. These include:

   a. Inherent risks associated with the handling of glassware, such as those used in this experiment.

   b. The slight risks involved in handling acidic and basic solutions. Prolonged contact with skin may cause minor irritation.

4. The following precautions will be taken during the experimental situation:

   a. The researcher will monitor your activities throughout the experiment. Warnings will be given if an unsafe situation occurs. If such situations persist, the experiment will be stopped. You will not be allowed to continue and will be paid for the amount of time spent up to that point.

   b. An eye-wash will be provided in the event of chemical contact with the eye.

   c. The solutions used have been diluted to lessen the risk of harm. Nevertheless, caution must be taken in their handling and use.

   d. A first aid kit will be located inside the room.

5. The data gathered in this experiment will be treated with anonymity. Shortly after the completion of the experiment, any connection between you and your data will be removed.

6. You should not volunteer for this study if you are not between the ages of 18-30 years. You should not volunteer if you are in poor emotional and/or physical health. It is your responsibility to inform the researcher of any other conditions that may interfere with your ability to perform the experiment.

7. You may ask the researcher any question you have about the study. Answers will be given to questions that do not compromise the validity of the study. Questions that may potentially bias the study will be answered in detail only after the second session is completed.

8. Data collection is expected to continue until the end of February. Please refrain from discussing this experiment with anyone until the beginning of March.

9. You are free to withdraw from the study at any time and for any reason. If you should decide to withdraw, you will receive payment for the time spent. You may also ask the experimenter to withdraw your data from analysis.

10. If you have any questions or concerns regarding this experiment, you may contact Dr. E. R. Stout, Chairman of the University's Institutional Review Board. His phone number is (703) 231-5284. The principle advisors to this experiment are Drs. Price and Geller. They can be reached at 231-5635 and 231-6223, respectively.

11. You should not sign this form until you are satisfied that you understand all of the conditions and descriptions stated above.

12. Your signature below indicates that you have read this document in its entirety and understand its content. Your signature also indicates that your questions have been sufficiently answered, and that you consent to participate in the study described. Include your printed name and address below is you would like a summary of the results to be sent to you.

______________________________ ______________________________

Signature

______________________________ ______________________________

Date

____

Printed Name and Address

APPENDIX B

Titration Answer Sheet

TITRATION ANSWER SHEET

Subject #: ______________ Date: ____________________

Age: ____________________ Session #: _______________

Trial #: __________________

Please follow the directions to the best of your ability. Record your data on this sheet only. Do not put any additional marks on this sheet.

1. Weight of Acid "A" (To the nearest 0.01 g): _______ _______ _______g

2. Vol. of Solution "S" (To the nearest 0.1 ml): _______ _______ _______ml

3. Volume of water (To the nearest 0.1 ml): _______ _______ _______ml

4. Drops of indicator (Please measure carefully): _______ _______ _______drps

5. Initial vol. of Base "B" (To the 0.1 ml): _______ _______ _______ml

6. Final vol. of Base "B" (To the 0.1 ml): _______ _______ _______ml

APPENDIX C

Laboratory Directions and Procedures

LABORATORY DIRECTIONS AND PROCEDURES

Purpose

The purpose of the experiment is to finish each trial accurately and in the shortest time safely possible.

Safety

Safety is a very serious matter. Please take the utmost care when working with all the equipment present.

The researcher will demonstrate safe handling procedures for the equipment involved. Feel free to ask questions after the demonstration has been completed. After the demonstration has been completed, you will be given 5 minutes to practice these handling procedures.

Both sessions in which you will be taking part will be videotaped. The researcher will also be observing the entire procedure while in the same room. If unsafe practices are observed, the researcher may either voice a warning or discontinue the experiment. In the event of the latter, payment will be given only for the time spent.

You will be asked to perform three trials of the same task during both sessions. For each trial, the equipment and chemicals will be identical. The directions will also be identical. After each session, you will be asked to answer a questionnaire pertaining to the trials.

Listed below are general safety procedures that must be followed:

1. If chemicals come in contact with the eyes, proceed to the eye wash station immediately and rinse eyes for a minimum of 15 minutes.
2. Full foot coverage is required at all times. Sandals, open-toed, and open-topped shoes are not permitted. Sneakers are permitted.
3. To minimize fire hazards, traffic problems, and accidents that may result from poor housekeeping, personal belongings will be kept in a safe area away from the work area.
4. Eating, drinking and smoking are not allowed. Chemicals are to be kept away from the mouth and eyes.
5. Long hair must be securely confined in a bun or a cap.
6. Old, comfortable clothing is recommended. Loose clothing, especially hanging sleeves, are not permitted.

APPARATUS AND MATERIAL

150ml Burets	Funnel
50ml Burets	Weighing Paper
250ml Erlenmeyer Flasks	Indicator
25 ml Graduated Cylinders	Acid "A"
Buret Clamp	Base "B"
Micro-Spatula	Solution "S"
Water	Recording paper

1. Erlenmeyer Flask: It can be distinguished by its conic shape.

2. Graduated cylinder: It is slender and has an red base.

3. Buret: It is held in a buret clamp. It is operated by turning the knob to the horizontal position to stop flow, and by turning the knob to the vertical position to allow flow.

4. Beaker: It has a wide opening and is cylindrical in shape. Please note the labelling on the beakers when using the solution it contains.

5. Weighing paper: The powdered acid that will be used for this experiment will be weighed while on this paper.

6. Scale: The scale is a digital readout scale that indicates weight up the the .01 grams.

7. Indicator: This is the indicator. It is in the brown bottle with the dropper lid.

ADDITIONAL INSTRUCTIONS

1. Use only the equipment provided for each solution. The equipment for a certain solution is labelled with a letter that represents that solution.

2. When pouring a solution into an empty container, either use the funnel provided or use a stirrer to guide the fluid into the container.

3. Handle the glass rods at the "black" ends only to minimize contact with the solutions.

4. When measuring solutions with a graduated cylinder, use the "meniscus" as the point of measurement. For this task, the meniscus is the lowest point of the convex surface of that solution. Your eyes must be at the level of the liquid's surface to obtain an accurate measurement.

5. Solution "B" produces vapors that may irritate your eyes. To minimize the release of these vapors, please cover its flask when not in use!

6. Acid "A" absorbs moisture in the air. Keep Acid "A's" container capped when not in use.

7. The weighing paper should be weighed on the scale and tared before the powdered acid is placed on the paper. Once the "Tare" button has been pressed, the scale should read "0.00". Only then should the powdered acid be placed on the weighing paper.

8. Wash your hands at the end of the trial if you come in contact with either the acid, base, or solution "S".

9. The Expected time of completion for each trial is **9** minutes. A timer will be placed near you so that you may keep track of expired time. If you have not finished the task by the end of 9 minutes, continue until you have completed the task.

10. Once you have poured out material from its original container, you may not return the residual amount back into the container. This pertains to everything except the water.

Objective: To determine how much of Base "B" is needed to neutralize Acid "A". This task should be performed with safety and accuracy always in mind.

- Recommended Time of Completion: 9 minutes.

Procedure:

1. Accurately measure 5 ml of solution "S", using the graduated cylinder labelled, "S". Use the meniscus of the solution as the point of measurement.

2. Place the graduated cylinder with solution "S" onto the digital scale. Record the reading of the scale and place solution "S" into the 250ml Erlenmeyer flask provided.

3. Using a sheet of weighing paper and the micro-spatula, weigh .50 grams of acid "A" on the digital scale. Record the amount actually used. Place the acid in the flask with solution "S"

Warning: the acid readily reacts with solution "S". Avoid inhaling the gas created by this reaction.

4. Measure 20 ml of water using the unlabeled graduated cylinder. Record the volume of water used. Add the water to the solution in the 250ml Erlenmeyer flask.

5. Gently swirl the solution until all the acid has dissolved.

6. Add two drops of the indicator to the solution. Accurately record the amount of drops actually used (Continue with the experiment even if more than two drops is accidentally added, but be sure to record this amount!!).

7. Fill buret with base "B" to the 40 ml line. Use the funnel provided to accomplish this task. Write down the initial volume of base "B" **to the nearest 0.1 ml**.

8. Titrate the acid solution against the base until the pink color of the indicator persists and is identical to the color of the neutralized model provided. **Swirl the flask continuously during titration to ensure mixing**. Record the final volume of base **to the nearest 0.1 ml**.

   Note: Because of the nature of the acid, expect a gradual color change to occur instead of a sudden one.

9. Promptly inform the researcher at completion of Step "8".

APPENDIX D

Session Quesionnaire

SESSION QUESIONNAIRE

Subject #: ______________ Date: __________________

Age: __________________ Session #: _____________

Please write down the number that best represents your opinion using the scale given below:

```
        1         2         3         4         5         6         7
        +---------+---------+---------+---------+---------+---------+
  Strongly Disagree                Neutral                  Strongly Agree
```

1. The task directions were clear and concise. ____________
2. The apparatus provided were easy to use. ____________
3. The task can be completed safely and accurately within the recommended time of completion. ____________
4. The safety precautions taken were quite adequate. ____________
5. I did not feel any anxiety while working in the laboratory environment. ____________
6. I did not feel fatigued at any point during the session. ____________
7. The safety equipment did not hamper my performance. ____________

Comments: ____

____

APPENDIX E

Perceived Risk Questionnaire
(Final Questionnaire)

FINAL QUESTIONNAIRE

Subject #: ______________ Date: __________________

Age: ___________________

Please write down the number that best represents your opinion using the scale given below each question:

1. Working without goggles and an apron caused me to _________ my pace.

```
 1          2          3          4          5          6          7
 +----------+----------+----------+----------+----------+----------+
Slow                         My Pace Was                      Quicken
                             Not Affected
```

2. I felt _________ when working without goggles and an apron.

```
 1          2          3          4          5          6          7
 +----------+----------+----------+----------+----------+----------+
Very Unsafe                   Somewhat                        Very Safe
                                Safe
```

3. Working with goggles and an apron made me feel _________.

```
 1          2          3          4          5          6          7
 +----------+----------+----------+----------+----------+----------+
Not much Safer               Opinion did                      Much Safer
                             Not Change
```

4. The solutions I worked with were _________ dangerous.

```
 1          2          3          4          5          6          7
 +----------+----------+----------+----------+----------+----------+
Very Dangerous                Somewhat                    Not Dangerous
                              Dangerous
```

5. Repeating the task during each session made me judge the chemicals as _________.

```
1           2           3           4           5           6           7
+-----------+-----------+-----------+-----------+-----------+-----------+
More Unsafe                     Opinion did                      Much safer
                                Not Change
```

6. I would judge my task performance pace as __________.

```
1           2           3           4           5           6           7
+-----------+-----------+-----------+-----------+-----------+-----------+
Slow and                        Fast but                         Quite Fast
Comfortable                     Safe
```

7. I was __________ concerned with accuracy while performing the task.

```
1           2           3           4           5           6           7
+-----------+-----------+-----------+-----------+-----------+-----------+
Not Concerned                   Somewhat                         Very
                                Concerned                        Concerned
```

8. I perceived __________ benefits working at the pace I had chosen.

```
1           2           3           4           5           6           7
+-----------+-----------+-----------+-----------+-----------+-----------+
No Benefits                     Some Benefits                    Many Benefits
```

9. I felt __________ while weighing the proper proportion of powdered acid.

```
1           2           3           4           5           6           7
+-----------+-----------+-----------+-----------+-----------+-----------+
Very Unsafe                     Somewhat                         Very Safe
                                Safe
```

10. I felt __________ while mixing the powdered acid with solution "S".

```
1           2           3           4           5           6           7
+-----------+-----------+-----------+-----------+-----------+-----------+
Very Unsafe                     Somewhat                         Very Safe
                                Safe
```

11. I felt __________ while mixing water with the dissolved acid solution.

```
1           2           3           4           5           6           7
+-----------+-----------+-----------+-----------+-----------+-----------+
Very Unsafe                     Somewhat                         Very Safe
                                Safe
```

12. I felt __________ while preparing the base for the titration.

```
      1          2          3          4          5          6          7
      +----------+----------+----------+----------+----------+----------+
Very Unsafe                        Somewhat                          Very Safe
                                     Safe
```

13. I felt __________ while performing the titration experiment.

```
      1          2          3          4          5          6          7
      +----------+----------+----------+----------+----------+----------+
Very Unsafe                        Somewhat                          Very Safe
                                     Safe
```

APPENDIX F

Zuckerman's Sensation Seeking Scale -- Form V

ZUCKERMAN'S SENSATION SEEKING SCALE -- FORM V

DIRECTIONS: Each of the items below contains two choices, A and B. Please circle the letter that most describes your likes or the way you feel. In some cases you may find items in which both choices describe your likes or feelings. Please choose the one which better describes your likes and feelings. In some cases you may find items in which you do not like either choice. In these cases mark the choice you dislike least. Do not leave any item blank.

It is important you respond to all items with only one choice, A or B. We are interested only in your likes or feelings, not in how others feel about these things or how one is supposed to feel. There are no right or wrong answers as in other kinds of tests. Be frank and give your honest appraisal of yourself.

1. A. I like "wild" uninhibited parties.
   B. I prefer quiet parties with good conversation

2. A. There are some movies I enjoy seeing a second or even a third time.
   B. I can't stand watching a movie that I've seen before.

3. A. I often wish I could be a mountain climber.
   B. I can't understand people who risk their necks climbing mountains.

4. A. I dislike all body odors.
   B. I like some of the earthy body smells.

5. A. I get bored seeing the same old faces.

B. I like the comfortable familiarity of everyday friends.

6. A. I like to explore a strange city or section of town by myself, even if it means getting lost.

B. I prefer a guide when I am in a place I don't know well.

7. A. I dislike people who do or say things just to shock or upset others.

B. When you can predict almost everything a person will do and say he or she must be a bore.

8. A. I usually don't enjoy a movie or play where I can predict what will happen in advance.

B. I don't mind watching a movie or play where I can predict what will happen in advance.

9. A. I have tried marijuana or would like to.

B. I would never smoke marijuana.

10. A. I would not like to try any drug which might produce strange and dangerous effects on me.

B. I would like to try some of the new drugs that produce hallucinations.

11. A. A sensible person avoids activities that are dangerous.

B. I sometimes like to do things that are a little frightening.

12. A. I dislike people who are free and easy about sex.

B. I enjoy the company of people who are free and easy about sex.

13. A. I find that stimulants make me uncomfortable.

B. I often like to get high (drinking liquor or smoking marijuana).

14. A. I like to try new foods that I have never tasted before.

B. I order the dishes with which I am familiar, so as to avoid disappointment and unpleasantness.

15. A. I enjoy looking at home movies or travel slides.

B. Looking at someones home movies or travel slides bore me tremendously.

16. A. I would like to take up the sport of water skiing.

B. I would not like to take up water skiing.

17. A. I would like to try surfboard riding

B. I would not like to try surf board riding.

18. A. I would like to take off on a trip with no preplanned or definite routes, or timetable.

B. When I go on a trip I like to plan my route and timetable fairly carefully.

19. A. I prefer the "down to earth" kinds of people as friends.

B. I would like to have friends in "radical" groups, such as artists or "punks."

20. A. I would not like to learn to fly an airplane.

B. I would like to learn to fly an airplane.

21. A. I prefer the surface of the water to the depths.

B. I would like to go scuba diving.

22. A. I do not mind having friends with nonconventional sexual preferences.

B. I stay away from anyone I suspect of not being heterosexual.

23. A. I would like to try parachute jumping.

B. I would never want to try jumping out of a plan with or without a parachute.

24. A. I prefer friends who are excitingly unpredictable.

B. I prefer friends who are reliable and predictable.

25. A. I am not interested in experience for its own sake.

B. I like to have new and exciting experiences and sensations even if they are a little frightening, unconventional, or illegal.

26. A. The essence of good art is in its clarity, symmetry of form and harmony of colors.

B. I often find beauty in the "clashing" colors and irregular forms of modern paintings.

27. A. I enjoy spending time in the familiar surroundings of home.

B. I get very restless if I have to stay around home for any length of time.

28. A. I like to dive off the high board.

B. I don't like the feeling I get standing on the high board (or I don't go near it at all).

29. A. I like to date members of the opposite sex who are physically exciting.

B. I like to date members of the opposite sex who share my values.

30. A. Heavy drinking usually ruins a party because some people get loud and boisterous.

B. Keeping the drinks full is the key to a good party.

31. A. The worst social sin is to be rude.

B. The worst social sin is to be a bore.

32. A. A person should have considerable sexual experience before marriage.

B. It's better if two married persons begin their sexual experience with each other.

33. A. Even if I had the money I would not care to associate with flighty rich persons like those in the "jet set".

B. I could conceive of myself seeking pleasures around the world with the "jet set".

34. A. I like people who are sharp and witty even if they do sometimes insult others.

B. I dislike people who have their fun at the expense of hurting the feelings of others.

35. A. There is altogether too much portrayal of sex in movies.

B. I enjoy watching many of the "sexy" scenes in movies.

36. A. I feel best after taking a couple of drinks.

B. something is wrong with people who need liquor to feel good.

37. A. People should dress according to some standard of taste, neatness, and style.

B. People should dress in individual ways even if the effects are sometimes strange.

38. A. Sailing long distances in small sailing crafts is foolhardy.

B. I would like to sail a long distance in a small but seaworthy sailing craft.

39. A. I have no patience with dull or boring persons.

B. I find something interesting in almost every person I talk to.

40. A. Skiing down a high mountain slope is a good way to end up on crutches.

B. I think I would enjoy the sensations of skiing very fast down a high mountain slope.

APPENDIX G

ANOVA TABLES

Task and Subtask Times, Accuracy, Task Errors, and Error Rate

TABLE 1

2x4 ANOVA SOURCE TABLE

Task Completion Time per Session

Source	df	SS	MS	F	p
Between					
Condition (C)	3	41.68	13.89	2.23	0.10
Subject (S)/C	32	198.94	6.22	-----	
Within					
Session (Ses)	1	91.58	91.58	90.59	0.0001
CxSes	3	4.26	1.42	1.40	0.26
Ses x S/C	32	32.35	1.01	-----	
Total	71				

TABLE 2

2x4 ANOVA SOURCE TABLE

Subtask 1 Times per Session

Source	df	SS	MS	F	p
Between					
Condition (C)	3	0.79	0.26	0.94	0.43
Subject (S)/C	32	8.97	0.28	-----	
Within					
Session (Ses)	1	2.99	2.99	34.19	0.0001
CxSes	3	0.04	0.01	0.14	0.94
Ses x S/C	32	2.80	0.09	-----	
Total	71				

TABLE 3

2x4 ANOVA SOURCE TABLE

Subtask 2 Times per Session

Source	df	SS	MS	F	p
<u>Between</u>					
Condition (C)	3	0.65	0.22	1.20	0.32
Subject (S)/C	32	5.79	0.18	-----	
<u>Within</u>					
Session (Ses)	1	1.85	1.85	25.47	0.0001
CxSes	3	0.32	0.11	1.46	0.24
Ses x S/C	32	2.32	0.07	-----	
Total	71				

TABLE 5

2x4 ANOVA SOURCE TABLE

Subtask 3 Times per Session

Source	df	SS	MS	F	p
Between					
Condition (C)	3	1.61	0.54	1.59	0.21
Subject (S)/C	32	10.79	0.34	-----	
Within					
Session (Ses)	1	2.31	2.31	29.37	0.0001
CxSes	3	0.54	0.18	2.26	0.10
Ses x S/C	32	2.52	0.08	-----	
Total	71				

TABLE 6

2x4 ANOVA SOURCE TABLE

Subtask 4 Times per Session

Source	df	SS	MS	F	p
Between					
Condition (C)	3	0.25	0.08	0.42	0.74
Subject (S)/C	32	6.35	0.20	-----	
Within					
Session (Ses)	1	0.74	0.74	27.16	0.0001
CxSes	3	0.07	0.02	0.83	0.48
Ses x S/C	32	0.88	0.03	-----	
Total	71				

TABLE 7

2x4 ANOVA SOURCE TABLE

Subtask 5 Times per Session

Source	df	SS	MS	F	p
Between					
Condition (C)	3	15.22	5.07	3.24	0.03
Subject (S)/C	32	50.09	1.57	-----	
Within					
Session (Ses)	1	14.05	14.05	35.37	0.0001
CxSes	3	0.52	0.17	0.44	0.73
Ses x S/C	32	12.72	0.40	-----	
Total	71				

TABLE 8

4x3 ANOVA SOURCE TABLE

Subtask 5 Times per Trial: Session 1

Source	df	SS	MS	F	p
Between					
Condition (C)	3	34.50	11.50	2.52	0.08
Subject (S)/C	32	145.79	4.56	-----	
Within					
Trial (T)	2	19.16	9.58	12.22	0.0001
CxT	6	2.14	0.36	0.46	0.84
T x S/C	64	50.16	0.78	-----	
Total	107				

TABLE 9

4x3 ANOVA SOURCE TABLE

Subtask 5 Times per Trial: Session 2

Source	df	SS	MS	F	p
Between					
Condition (C)	3	15.66	5.22	4.36	0.01
Subject (S)/C	32	38.31	1.20	-----	
Within					
Trial (T)	2	0.27	0.14	1.95	0.15
CxT	6	0.83	0.14	1.99	0.08
T x S/C	64	4.46	0.07	-----	
Total	107				

TABLE 10

2x4 ANOVA SOURCE TABLE

Errors per Session

Source	df	SS	MS	F	p
Between					
Condition (C)	3	1.20	0.40	1.74	0.18
Subject (S)/C	32	7.38	0.23	-----	
Within					
Session (Ses)	1	3.27	3.27	14.44	0.0006
CxSes	3	1.94	0.65	2.87	0.05
Ses x S/C	32	7.23	0.23	-----	
Total	71				

TABLE 11

2x4 ANOVA SOURCE TABLE

Error Rate per Session

Source	df	SS	MS	F	p
Between					
Condition (C)	3	0.05	0.02	3.20	0.04
Subject (S)/C	32	0.16	0.005	-----	
Within					
Session (Ses)	1	0.01	0.01	2.29	0.14
CxSes	3	0.04	0.01	2.77	0.06
Ses x S/C	32	0.14	0.004	-----	
Total	71				

TABLE 12

2x4 ANOVA SOURCE TABLE

Subtask 1 Accuracy per Session

Source	df	SS	MS	F	p
Between					
Condition (C)	3	0.07	0.2	1.25	0.31
Subject (S)/C	32	0.58	0.02	-----	
Within					
Session (Ses)	1	0.004	0.004	0.27	0.61
CxSes	3	0.06	0.02	1.28	0.30
Ses x S/C	32	0.46	0.01	-----	
Total	71				

APPENDIX H

Session Questionnaire ANOVA

TABLE 1

ANOVA SOURCE TABLE

First and Second Session Questionnaire Difference Score: Item 1

Source	df	SS	MS	F	p
Between					
Condition (C)	3	0.44	0.15	0.67	0.58
Subject (S)/C	32	7.11	0.22	-----	
Total	35				

TABLE 2

ANOVA SOURCE TABLE

First and Second Session Questionnaire Difference Score: Item 2

Source	df	SS	MS	F	p
Between					
Condition (C)	3	0.67	0.22	0.23	0.87
Subject (S)/C	32	30.88	0.97	-----	
Total	35				

TABLE 3

ANOVA SOURCE TABLE

First and Second Session Questionnaire Difference Score: Item 3

Source	df	SS	MS	F	p
Between					
Condition (C)	3	4.97	1.66	2.41	0.09
Subject (S)/C	32	22.00	0.69	-----	
Total	35				

TABLE 4

ANOVA SOURCE TABLE

First and Second Session Questionnaire Difference Score: Item 4

Source	df	SS	MS	F	p
Between					
Condition (C)	3	54.97	18.32	9.46	0.0001
Subject (S)/C	32	62.00	1.94	-----	
Total	35				

TABLE 5

ANOVA SOURCE TABLE

First and Second Session Questionnaire Difference Score: Item 5

Source	df	SS	MS	F	p
Between					
Condition (C)	3	4.75	1.58	0.65	0.59
Subject (S)/C	32	78.00	2.43	-----	
Total	35				

TABLE 6

ANOVA SOURCE TABLE

First and Second Session Questionnaire Difference Score: Item 6

Source	df	SS	MS	F	p
Between					
Condition (C)	3	0.08	0.03	0.02	0.99
Subject (S)/C	32	37.56	1.17	-----	
Total	35				

APPENDIX I

Final Questionnaire ANOVA

Item 6

Source	df	SS	MS	F	p
Between					
Ses. 1 Condition (C1)	1	7.11	7.11	5.60	0.02
Ses. 2 Condition (C2)	1	0.11	0.11	0.09	0.77
C1xC2	1	0.00	0.00	0.00	1.00
Subject (S)/ C1xC2	32	40.67	1.27	——	
Total	35				

APPENDIX J

t-Statistic Equations

t-Statistic: use of this formula depends on the assumption that the population variances of the two groups are equal.

$$t = (x_1 - x_2) / [s^2 (1/n_1 + 1/n_2)]^{1/2}$$

where s^2 is the pooled variance

$$s^2 = [(n^1 - 1) s_1^2 + (n^2 - 1) s_2^2] / (n^1 + n^2 - 2)$$

$$df = n_1 + n_2 - 2$$

Separate variance t-test: used under the assumption of unequal population variances.

$$t' = (x_1 - x_2) / (w_1 + w_2)^{1/2}$$

where $w = s_1^2 / n_1 \;,\; s_2^2 / n_2$

$$df = (w_1 + w_2)^2 / [w^2 / (n_1 - 1) + w^2 / (n_2 - 1)]$$

VITAE

Rafael Kent Angeles Marshall was born in Quezon City, Philippines on June 9, 1966. He received his Bachelor of Art degree in psychology from George Washington University in 1988. Rafael was employed at Carlow Associates, Inc., as a human factors research assistant for one year before pursuing his Master of Science degree from Virginia Polytechnic Institute and State University. Rafael is an affiliate of both the American Society of Safety Engineers (ASSE) and the Human Factors Society (HFS).