5-2013

An Experimental Study on the Influence that Failure Number, Specialization, and Domain have on Confidence in Predicting System Failures

Somaiah Thimmaiah
Clemson University, sthimma@clemson.edu

Recommended Citation

Thimmaiah, Somaiah, "An Experimental Study on the Influence that Failure Number, Specialization, and Domain have on Confidence in Predicting System Failures" (2013). *All Theses*. 1654.
https://tigerprints.clemson.edu/all_theses/1654

AN EXPERIMENTAL STUDY ON THE INFLUENCE THAT FAILURE NUMBER, SPECIALIZATION, AND DOMAIN HAVE ON CONFIDENCE IN PREDICTING SYSTEM FAILURES

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Mechanical Engineering

---

by
Somaiah Muccatira Thimmaiah
May 2013

Accepted by:
Dr. Joshua D. Summers, Committee Chair
Dr. Fred S. Switzer
Dr. Georges M. Fadel

ABSTRACT

Design reviews are typically used for three types of design activities: 1) identifying errors, 2) assessing the impact of the errors, and 3) suggesting solutions for the errors. This experimental study focuses on understanding the second issue as it relates to the number of errors considered, the existence of controls, and the level of domain familiarity of the assessor. A set of design failures and associated controls developed for a completed industry sponsored project is used as the experimental design problem. Non-domain individuals (psychology class students), domain generalists (first year engineering students), and domain specialists (graduate mechanical students) are provided a set of failure modes and asked to estimate the likelihood that the system would still successfully achieve the stated objectives. Primary results from the study include the following: the confidence level for all domain population decreased significantly as the number of design errors increased (largest p-value=0.0793) and this decrease in confidence is more significant as the design errors increase. The impact on confidence is less when solutions (controls) are provided to prevent the errors (largest p-value=0.0334), the confidence decreased faster for domain general engineers as compared to domain specialists (p=<0.0001). The domain specialists showed higher confidence in making decisions than domain generalists and non-domain generalists as the design errors increase.

The research presents a study on how estimations are made in design reviews. It answers the question on how individuals assess the performance of systems which is necessary to be addressed in order to evaluate the importance of methods such as design reviews and design review tools (FMEA, DFMEA, FTA) used in design engineering. It addresses the challenges faced by the impact of design errors in the design process and how they affect assessment by different types of designers in predicting successful system performance.

DEDICATION

To my dad, Thimmaiah and mom, Manjula, for their continuous support towards my graduate studies

To my uncle, Kusha and aunt, Ashique for their support during my time in Clemson University

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Joshua Summers, for the continuous guidance and financial support he offered throughout the course of this research.

I thank Drs. James Rieck, Fred Switzer, Georges Fadel and Scott Shappel for their assistance in this research, and CEDAR group members for their support.

I thank the student participants from Clemson University, the departments of General Engineering, Mechanical Engineering, and Psychology for supporting this work.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (CONTINUED)

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER ONE: MOTIVATION

During my final year as an undergraduate, I worked on a project titled "Electronically Controlled Infinitely Variable Pressure Control for Chuck or Tail Stock of CNC Lathe'. This project, supported by Bosch Rexroth, dealt with the design and fabrication of a hydraulic power pack. The hydraulic system can be connected to the headstock or tailstock of a CNC Lathe and the pressure can be varied electrically to achieve the required clamping force of a hydraulic chuck. The variation of clamping force for shafts of different materials can be achieved through this. The project was completed by a team of four undergraduate engineering students, three shop floor maintenance personnel, one senior engineer with expertise in sales of hydraulic products, and two senior hydraulic engineers. The duration of the project was 3 months.

Over the course of the project, several issues had to be addressed. As an example, the hydraulic system required a working pressure of 10 bars to achieve the required clamping, however a maximum pressure of only 6-7 bars could be achieved with the current resources. The project team, comprised of individuals with varying years of experience had frequent discussions on troubleshooting in order to achieve the required pressure. During the design reviews, assessing the criticality of the design issues was our primary target. Reflecting on the reviews, the design review team was more reliant on the remarks made by the experienced individuals regarding how the problem could be fixed. More importantly the remarks made by individuals who had the most experience in the field of hydraulics were regarded more highly since the troubleshooting process *began with the remarks made by these individuals. Eventually the problem was fixed by replacing the pressure relief valve with minor adjustments and this solution was suggested by one of the domain experts.*

In industry during design review, engineers gather to discuss how important the design errors or the failure modes are in terms of criticality. The design review team's expertise may be distributed across different domain population in the organization and the time spent by the domain experts is expensive. From my experience on this project, the team members often did not agree while assessing the importance of failure modes. The design team preferred assessing failure modes individually with the most critical failure mode addressed first. This motivated me to wonder if considering multiple failure modes at a given time, and the level of domain awareness or knowledge had an effect during failure mode assessment.

A systematic collaborative method is necessary to overcome challenges such as considering multiple failure modes at a given time and/or considering the domain awareness of the individuals present during design review to make decisions with confidence. The main goal of this research is to develop a method to calibrate the assessment of failure modes by all individuals in the design team.

1.1 Understanding How Designers Assess Possible Failure

Companies seek to manufacture products with low production failures and high use reliability without major cost increases. Improving a design is often simpler if the reliability and risk of a product is assessed early in the product development process (R. Schmidt, 2010). The cost and time spent on a failure detected during production is high compared to a failure detected during the product development process. It has been shown that costs can increase by a factor of ten with each subsequent phase (Figure 1) during the product development process (Pfeiffer, 2002). Thus, it would be useful to spend time and effort towards risk assessment at an early stage and this will help to save costs and time during subsequent product development projects. Failure to detect design problems early in the design process can become evident when product problems arise after the product has been put into production. As an example, the Ford Motor Company had to recall their vehicles when it was discovered that vehicles equipped with adjustable pedals being positioned too close, caused drivers to unintentionally hit the accelerator when trying to slow down[1]. For a good risk assessment during the product development process, structured methods such as Failure Mode and Effect Analysis (FMEA) (Mach & Duraj, 2008), Design Review Based on Failure Mode (DRMFM) (Stamatis, 2003), or Fault Tree Analysis (FTA) (R. Schmidt & Spindler, 2012) can be included.

---

[1] Ford recalling Tauruses & Sables because pedals are too close – 2002 - http://www.firstcoastnew.com/onyourside/articles/2002-10-09/recalls_ford.asp

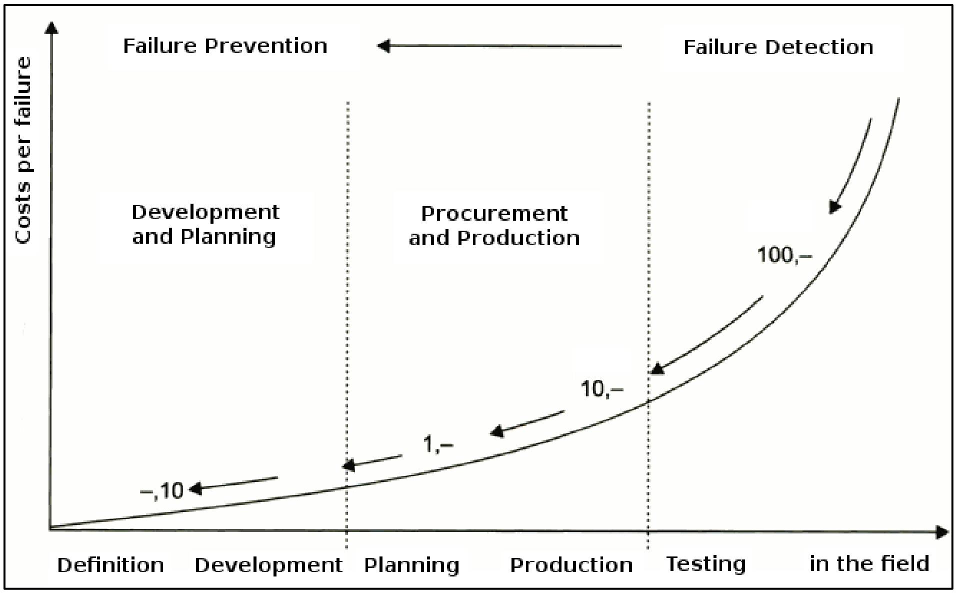


Figure 1: Increase of costs per failure related to a failure detected in different parts of the production cycle (Pfeiffer, 2002)

The goal of achieving low failure rates and high reliability in products cannot be achieved by attributive quality control or process control only (R. Schmidt, 2010). Quality cannot be achieved through testing and improving a product, it has to be built in from the beginning of the design process and maintained throughout the production process (Pahl, Beitz, Wallace, & Lucienne Blessing, 2007; Peace, 1993). Ensuring quality and improving quality are team activities that can be achieved during design review. Quality is influenced decisively during design and development and is realized during production. Nissan uses the design review method to develop higher-quality parts[2]. During the design review, design experts work together to assess the potential risk for each part, and devise ways to prevent problems proactively. Design review

---

[2] Initiatives by NISSAN Quality – 2009, http://www.nissan-global.com/EN/QUALITY/PRODUCTS

conducted by certified personnel allows quicker and more accurate design inspections, and aids the development of problem-free parts.

Managing the decisions made by a team of collaborating experts becomes challenging with the increasing demand for complex and interrelated systems. This is especially challenging during the early stages of product development since there is limited knowledge, high uncertainties, and the decisions made have far reaching effects on the directions pursued thereafter, and hence the affordability, reliability/safety and effectiveness of the final product (Ullman & Spiegel, 2006). In early phases of the design, much of the information is qualitative, rendering the early decisions subjective (Ullman & Spiegel, 2006). However, efforts towards making good decisions at this stage have high payoffs. While designing a product, such as a tire, the designer should ask whether the crux is to improve the technical functions, such as traction through tread structures, to lower the cost, to shorten the delivery times, or to improve the production methods. All of these mentioned requirements while designing a product should be satisfied but their importance might not be equal.

Identifying the requirements as critical can be subjective during the early stages of the product development. Understanding the subjectivity and the confidence that engineers impart on decision making during the assessment of the impact of multiple failure modes or design errors on the performance of the system is the focus of this research.

1.2 Path of Thesis

Chapter Two presents an understanding of design reviews; its role in achieving stakeholders' requirements; conducting design reviews along the design process to identify technical risks in performance, manufacturing, testing, and use; steps involved in conducting design review; challenges and issues faced during design review. In addition, Chapter Two presents the role played by design reviews in achieving quality. Specifically, the design review tool (FMEA) used in this research and scales to measure confidence are discussed.

Chapter One presents the experimental user study developed to investigate the impact of the number of design errors on the assessment of system success for decision-making during a design review. The design problem for review is the tent ballast testing equipment. A study involving 143 general engineers, 43 psychology students, and 25 graduate students is deployed with the data collection based on linear scale markings.

Chapter Four presents the analysis and results of the user study. The confidence ratings are used to compare the impact of design errors on confidence. Findings on the performance of domain specialists, non-domain generalists, and domain generalists for interaction between domains and design errors (with and without controls) are discussed.

Chapter Five concludes the research and an understanding is derived that relates to improving the performance and confidence during design review. Areas that require further investigation are also discussed.

1.3 Key Findings from the Research

The domain specialists (graduate mechanical engineering students) showed higher confidence in making decisions than domain (freshmen engineering students) and non-domain generalists (psychology students) for increasing number of design errors. The difference in confidence between the three domains is more evident as the design errors increase and the difference is greater as the number of errors increase. Higher confidence of the domain specialists suggests that domain specialists were more optimistic in the system success (higher confidence) while making decisions with design errors. When a single design error is considered at a time, the domain knowledge of the individual does not matter as the domain population can be considered equivalent. However, while considering more than one design error, there is a significant difference in the confidence of the decision. This suggests that decision making on multiple design errors should be carefully considered since the domain experience begins to be an issue. The confidence level for all populations decreased significantly as the number of design errors increased. For a maximum of seven design errors, confidence level below 30% was attained for domain generalists and non-domain generalists. Design errors beyond seven are considered insignificant since the confidence level is too low to make high value decisions. The confidence level increases when solutions (controls) are provided to mitigate the errors. For example, the confidence level for three design errors increased from 50% to 61% with controls ($p=0.0034$). Hence, this enables engineers to consider more errors at a given time while using controls. The domain specialists showed higher confidence in predicting the performance of the system as the design errors increase.

CHAPTER TWO: DESIGN REVIEWS

Design reviews play an important role in ensuring that the product or design artifact meets the requirements of various stakeholders (Arthur & Groner, 2004; Hisarciklilar & Boujut, 2009; Mohan, Jain, & Ramesh, 2007). These reviews are critical in reducing risks by identifying problems, assessing the impact of the problems, and suggesting solutions to prevent the problem (Ostergaard, Wetmore III, Divekar, Vitali, & Summers, 2005; Sater-Black, K. and Iverson, 1994; Wetmore III, Summers, & Greenstein, 2010; Wetmore III & Summers, 2003). Typically, a design review is a collaborative, synchronous meeting where the current state of the solution is compared against the defined problem. The participants might include engineers, manufactures, marketers, and suppliers. The objectives are to identify potential flaws or challenges in the solution, assess the impact that these flaws have on the overall performance, prioritize the resolution of these errors, and possibly offer new controls to mitigate them.

Design review is critical in industry to reduce risk in design projects and can provide the necessary discipline and methodology for timely identification of design problems and their solutions (Chapman, 1998; Clarkson & Eckert, 2004; C. Schmidt, Dart, Johnston, Sterling, & Thorne, 1999; Wetmore, 2004). Often, these reviews are conducted several times throughout the design process (conceptual design, embodiment design, detail design) to identify technical risks in performance, manufacturing, testing, and use (Collins, Yassine, & Borgatti, 2009). Design reviews are used iteratively and are dependent on the different stages of development, the legal requirements, the industry best practices, and the company culture. The reviews may include inputs from individuals with expertise in various domain population such as customer, development team, or suppliers (Ostergaard et al., 2005). Input from several functional groups develops a more accurate view of the design artifact, thereby improving the likelihood that failures can be identified (Wetmore III et al., 2010). This is due to the larger information and expertise base, influence of superior decision-making from interacting groups, and the checking of errors and rejection of flawed suggestions (Hammond, Lafayette, Koubek, & Harvey, 2001).

The first step in conducting a design review is to identify the participating individuals by listing the characteristics of the design and identifying the resources needed for the characteristics to be discussed (Pugh, 1991). It is necessary for individuals with expertise to be present during design review. For example, if a hydraulic power unit is being reviewed for functionality, then individuals with expertise in areas such as sales and ergonomics might be lower in priority for inclusion on the review team. During the design review, evaluation is conducted to highlight potential deficiencies that reduce the design performance as viewed by various stakeholders. The design review tools discussed later enhance product design by improving design performance and quality.

Quality cannot be achieved through testing and improving a product, it has to be built into the product from the beginning of the design process and maintained throughout the production process (Pahl et al., 2007). Ensuring quality and improving quality are team activities that can be achieved during design review. Quality is influenced decisively during design and development and is realized during production.

Nissan uses the design review method to develop higher-quality parts. During the design review, design experts work together to assess the potential risk for each part, and devise ways to prevent problems proactively (“Initiatives-NISSAN, Quality,” 2009). Design review conducted by certified personnel allows quicker and more accurate design inspections, and aids the development of problem-free parts. By predicting the performance of the system early in the product development stage can eliminate potential problems in the final product.

The stage during which a design review is conducted may depend on the product, company, or team members, it is an iterative process in which earlier stages of the design process are revisited, and the design is altered, reflecting changes to eliminate the identified problem (Wetmore, 2004). The required expertise represented by the identified individuals may be distributed across different domain population in the organization. Ideally, domain specialists with significant experience should be included, but these individuals’ times are costly. Consequently, design review teams are facing new challenges with task efficiency and effectiveness of high-level decision-making (Hilts, S., Johnson, K. and Turoff, 1986). To address these challenges, a systematic collaborative method is needed to overcome the problems faced during decision making encountered by involving individuals of expertise in various domain population. By replicating the method used in this study, individuals with the expertise in associated domain can be correctly identified for the review or the assessments made by non-domain specialists can be calibrated for similar assessments as those of domain specialists.

2.1 A Design Review Tool: Failure Modes and Effects Analysis (FMEA)

The Failure Modes and Effects Analysis (FMEA) tool was originally developed to improve the reliability of complex systems in 1949 by the US Military as the Failure Mode Effect and Criticality Analysis (Department of Defense, 1980). It is one of the most popular tools to virtually test the reliability of a product, process, or a system in many different industries ranging from automotive to aerospace to engineered-to-order equipment (Teoh & Case, 2005). It requires knowledge on how a system and its components are susceptible to failure in order to assist engineering to deliver a reliable product (G. Hawkins & Woollons, 1998; Teoh & Case, 2005; Xu, Tang, Xie, Ho, & Zhu, 2001). The output information from FMEA can be used to guide the design and redesign process to focus on critical areas. It helps analyzing risk which has a great potential of cost savings as potential problem areas are identified and corrected (Cotnareanu, 1999; Yasenchak, 2000). It is recommended during FMEA to include a team of knowledgeable individuals as all aspects of the product are evaluated[3]. The team can involve individuals from the design team, manufacturing team and suppliers.

The standard process involves exploring the entire system and identifying potential failure modes, determining the effect of those failures, and how critical these failures effects are with respect to product functionality (Teng, 1996). A list of the potential failure modes for each of the functions of the system is listed. The consequence of each failure mode on the system's performance is evaluated. The root causes of each failure mode are evaluated against the activities for prevention and detection. The final

---

[3] Potential Failure Mode and Effects Analysis Reference Manual

step is to score the complete form with severity, occurrence, and likelihood of detection, all in a scale from 1-10. Subjective values ranging from 1-10 are defined for the respective (severity, occurrence, detection) columns and the three scores are then multiplied to obtain the risk priority number (RPN). Higher values indicate immediate corrective action. Additional activities, or design changes to reduce the probability of failure are added and the RPN changes with these improvements. The definitions of the terms used in FMEA are mentioned below (Chrysler Corporation, 1995)

- Potential failure mode: It is the manner in which the component or subsystem potentially fails to design intent (ex: break, leak)
- Effects: It is the consequence of each failure mode (ex: complete failure, pressure not maintained)
- Severity: It is the factor that represents the seriousness or impact of the failure to the customer or to a subsequent process. Severity is ranked from 1 to 10 with 1 being least severe and 10 being most severe
- Occurrence: It is the likelihood that a specific failure mode will occur. Occurrence is ranked from 1 to 10 with 1 being unlikely failure and 10 being highly likely
- Detection: It is the ability of the design controls to detect a potential failure mode before it leaves the facility. Detection is ranked from 1 to 10 with 1 being almost certain to detect a potential cause and 10 being the control will not detect a potential cause

- Risk Priority Number (RPN): It is a measure used in assessing risk to help identify critical failure modes associated with the design or process. RPN can vary from 1 (absolute best) to 1000 (absolute worst)

The primary method used in FMEA to identify failure modes is brainstorming (Davis, Stanley, William Riley, Ayse P. Gurses, Kristi Miller, 2008). The RPN are normally generated from expert opinion and statistical estimates. The weakness of this is that assessment of potential risks and their underlying causes is based solely on domain expert's memory and knowledge (Bonnabry, Despont-Gros, & Grauser, 2008). Table 1 illustrates an example FMEA excerpt from a project to design a pressurized mud box for a shaft-seal manufacturer.

Table 1: Example FMEA Worksheet for a Pressurized Mud Box Seal Testing System

Function or Item	**Potential Failure**	**Potential Reason for Failure**	**Effects of Potential Failure**	**O**	**S**	**D**	**Score (RPN)**	**Action**
Shaft	Break	Fatigue	Complete Failure	1	10	10	100	Oversize Shaft
		Yielding	Complete Failure	1	10	3	30	None
Seal	Leak	Bolts elongate	Pressure not maintained	3	3	1	9	None
------	-----	-----	------	----	----	----	----	-----

O: Chance for occurrence; S: Severity; D: Chance for Detection

2.2 Measuring Confidence

The focus in this section is on scales to measure confidence and how the scale can be used to calibrate the confidence level and the number of failure modes. Calibration plays a critical role in individuals' ability to successfully self-regulate their own learning (Dinsmore & Parkinson, 2012) and their self-critique of their decisions. It is this self-critique, or the ability to understand the tendency of different groups of people to overestimate or underestimate the likelihood of failure of a system, that is of interest in this research. It is critical to make valid conclusions about the measurement of confidence as the failure modes increase and the calculations of the calibration used. Currently, there appears to be little consensus on what methods should be used to calculate this calibration (Dinsmore & Parkinson, 2012).

Dichotomous or categorical measures of confidence can be problematic since there is a possibility that people tend to choose any of the variables rather than adhering to a true dichotomy (Thorndike, E. L., & Gates, 1929). Complex interactions between the human, their behavior, and the environment suggest more complicated judgments than simply "confident" or "not confident". This further suggests that individuals' may be considering multiple criteria when making their judgments. Dichotomization of individuals' confidence could result in information loss leading to poor sensitivity analysis and an increase in the likelihood of "Type I errors" (Pedhazur, 1997). A type 1 error occurs when the null hypothesis is true, but is rejected. It is a focus of skepticism and occurs when we believe a falsehood (Shermer, 2002). For example, in the interaction between confidence level and number of design error, the null hypothesis is: The number of design errors does not affect the assessment of system performance. Type I error in this example is rejecting the null hypothesis when it is true. Using Likert-type scales with seven or more categories may overcome some of these issues when used in a structural equation model (Finney, S.J. & DiStefano, 2006).

A more robust approach might be to use the “100-mm line”, a data collection technique that has been widely used in educational psychology literature to measure various learning concepts (Schraw, G., Potenza, M. T., & Nebelsick-Gullet, 1993). It enables the collection of more precise information as the participant gives their measure of confidence across a sliding scale, as opposed to just “no confidence or full confidence”. The disadvantage in using the 100-mm scale is the interpretation of precisely how each participant chose to mark the line (Dinsmore & Parkinson, 2012). If a participant marks their response at 35mm and another participant marks at 55 mm, there is difficulty in explaining the difference in the measurement. Therefore a large sample size is used and the average of all participant confidence measurements is used to derive inferences. It is this latter approach that will be used in this experiment.

CHAPTER THREE: THE EXPERIMENT

3.1 Experimental Study

An experimental user study is developed to investigate the impact of the number of design errors on the assessment of system success for decision-making during a design review. The dependent variable of interest is the assessment of success or the confidence level as determined by individuals. There are three independent variables studied. The first, and primary, is the number of design errors present (one design error, three design errors, five design errors, and seven design errors). This independent variable (design error) is varied by increasing the number of design errors, with the impact on the dependent variable (confidence) being evaluated. The secondary independent variables studied include the condition of the error; is it presented with or without proposed mitigating controls. The third variable of interest is the type of individual doing the assessment. These variables are used in designing experiments to assess specific influences in a controlled environment that simulates portions of real-world design activities.

Table 2 provides a summary of the experiment layout. The layout is structured for participants from three different backgrounds (domain generalists = freshmen general engineering students; domain specialists = graduate mechanical engineering students; and non-domain generalists = junior psychology students). Each population is divided into two experimental groups, where one group is presented design errors without any controls and the second group is presented design errors with a proposed set of controls.

Table 2: Experimental Design

Scenario	**Design Errors Presented to Students**
Scenario 1	**1 design error:** Error A
Scenario 2	**3 design errors:** Error B Error C Error A
Scenario 3	**5 design errors:** Error D Error E Error B Error C Error A
Scenario 4	**7 design errors;** Error F Error G Error D Error E Error B Error C Error A

Population	**Controls**	**Scenario**			
Domain Generalist Student A	With	1	2	3	4
Domain Generalist Student B	Without	1	2	3	4
Domain Specialist Student A	With	1	2	3	4
Domain Specialist Student B	Without	1	2	3	4
Non-Domain Generalist Student A	With	1	2	3	4
Non-Domain Generalist Student B	Without	1	2	3	4

The author hypothesizes that the participants' confidence in the tent testing mechanism will decrease as the number of presented design errors increase. This decrease in confidence will be smaller when the design errors are presented with their associated controls. It is expected that the change between the confidence levels for the controlled and uncontrolled design errors will be greatest for the students with an engineering specialization. This belief is due to the engineering students' better understanding of the causes behind the design errors, affording them a better appreciation of how the controls will limit the effects of the design errors.

3.2 Experiment Problem: Tent Testing

As a concrete motivating example, an industry sponsored project to support framed-tent ballast performance testing executed by a team of eight graduate mechanical engineering students in a three month project at Clemson University is considered. In this project, test equipment and a testing protocol were developed to collect data on the movement resistance (friction) of different ballast types (concrete barrels, water barrels, cement blocks), different surface conditions (dry pavement, wet-smooth concrete, grass, gravel), with different modifying interfaces (neoprene, plywood, steel plates). These resistance coefficients are used in an industry tool to help large tent manufacturers, installers, and renters to determine the appropriate configuration of ballast needed.

In this project, risk assessment was done informally throughout the project in weekly meetings as the project team developed, prototyped, and built the unique testing equipment. Several possible failures were identified and corrective measures implemented, through these design review sessions. These design reviews were conducted to assure that the input requirements from the sponsor were being met. Reviews were conducted before and after build, and before and after the first test to ensure, as best possible, conformance to the requirements. Further, failures that were encountered in the testing equipment and testing procedure were corrected as the design progressed. An integrated FMEA worksheet was assembled to capture the predicted and the actual failures and their corrective actions. This integrated worksheet was reviewed by all team members for accuracy.

3.3 Tent Ballast Testing (Equipment and Process)

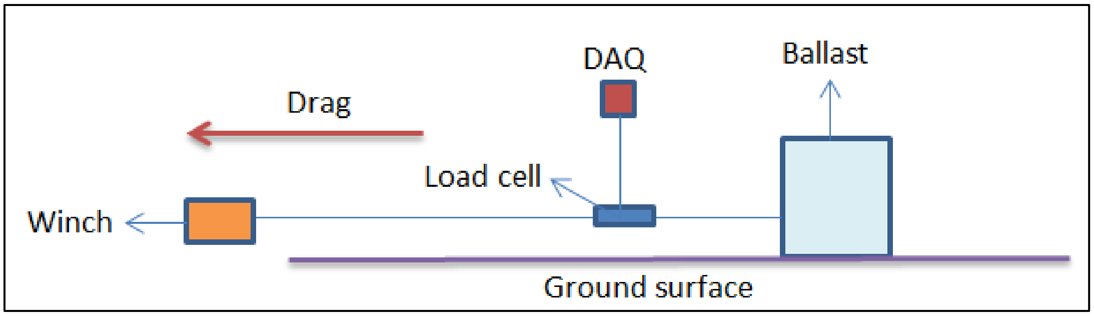
The design problem for review was the tent ballast testing equipment and process that was developed, implemented, and deployed for an industrial sponsor (IFAI–Industrial Fabric Association International). This project was selected because of the accessibility for individuals to conceptualize the problem as evidenced by the inclusion of several undergraduates on the initial project and by the fact that the project was of a short duration with full design, build, and use spanning approximately three months. Moreover, comprehensive, including design review results, testing results and fabrication/design changes was available. Finally, the project provided a system of low complexity that included failure potentials associated with the physical embodiment, the software monitoring, and the human test execution. This variety was sought to provide variety of error types to the experimental population.

The goal of the IFAI project was to develop systematic guidelines for selecting and configuring ballasts (weights) to tents when they cannot be staked to various ground conditions. In order to accomplish this, the coefficients of friction between the different types of ballast and ground surface are determined experimentally. The coefficient of friction values obtained are be used as a factor to aid in selecting the type and the number of ballasts for a tent on a particular type of ground surface.

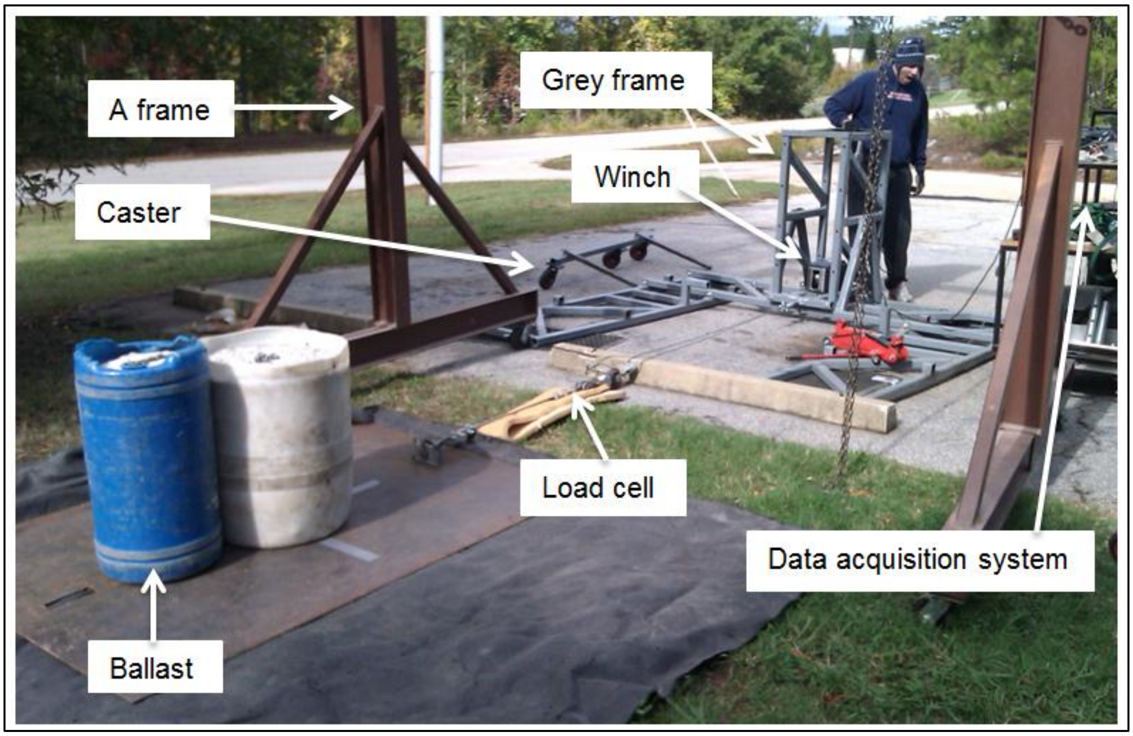
3.3.1 Testing Setup

Figure 2 shows the setup for the tent ballast drag test. The setup consists of a grey frame structure on which the winch is mounted. The front end of the load cell is attached to the winch rope and the rear end is hooked to the ballast. Data from the load cell is collected using a data acquisition module.

The testing was completed by eight mechanical engineering students, graduate and undergraduate, and one material science and engineering student. The friction coefficients between ballasts and different ground surfaces, including various modifiers, were obtained and the overall result of the project was lauded by the project sponsors. The collected data has been used to inform an on-line ballast configuration tool that has been deployed in the industry.



(a) Iconic Schematic of Test Setup



(b) Implemented Schematic of Test Setup

Figure 2: Tent Ballast Testing (Equipment) Setup

3.3.2 Testing Procedure

In addition to the test equipment setup, the testing procedure was defined as part of the project. This procedure is also provided as part of the experimental problem to the study participants. The test conducted consists of the following steps:

- Move grey frame to testing location
- Move ballast to testing location using A-frame
- Connect front end of the load cell to the winch rope
- Connect rear end of load cell to the ballast
- Attach data acquisition module cable to the load cell
- Operate winch until the ballast movement
- Stop winch after ballast movement
- Collect the data and calculate friction coefficient

After the completion of the design, build, and implementation of the project, an external FMEA was conducted to determine failure modes in the designed system. These failure modes were compared with those identified from project design review notes. Further, the failure modes and their ultimate implemented controls were verified through interviews with the project participants. As a result, sixteen failure modes, or design errors, were identified.

Table 3: Sixteen Identified Design Errors

#	**Identified Design Errors**
1	Releasing caster from grey frame using a hydraulic jack
2	Swaying of bars in the A frame while moving to the test location
3	Swaying of ballast in the A frame while moving to the test location
4	Connection failure between the data acquisition system and load cell
5	Breakage of winch cable
6	Control of winch by the operator during drag test
7	Collection of surface material below the ballast during a drag test
8	Irregular ground surface during drag test
9	Insertion of hook from the load cell to the ballast
10	Failing to weigh the ballast at the end of the test procedure
11	Failing to calibrate the load cell
12	Use of winch by different operators
13	Positioning of load cell during calibration
14	Instability in grey frame during drag test
15	Rope which prevents ballast from swaying breaks
16	Failing to calibrate the data acquisition system

In this manner, the errors and experimental problem that are presented to the experiment participants are derived from reality, thereby increasing the relevance to the students. This sense of realism and relevance is important to experimental problem selection to help increase the likelihood of student engagement in the experiment.

3.3.3 Design Error Pruning and Selection

The sixteen failures identified are known to be actual failures that were addressed in the project. However, it is not known whether these sixteen errors are perceived as equivalent. If one failure is considered much more significant by the experimental participants, then this would bias the assessment of the impact of the failures. Therefore, before conducting the experiment, it is necessary to conduct an initial study to determine which design errors are equivalent and can be used in the full experiment. In order to provide meaningful results, the design errors must all result in a similar assessment of impact on performance and must be independent of one another. Otherwise, the experiment would result in increased changes in confidence due to the type of design error rather than the number of design errors.

In order to prune the available design errors, 29 participants (twenty senior undergraduates and nine graduate level engineering students) within a design for manufacturing class were each randomly given five individual design errors out of the sixteen available. The design errors used in the pilot study did not include controls. It is important to determine the perceived or assessed impact due to the possible design error without any correction. The participants were instructed to provide their confidence in the success of the test ballast testing equipment and procedure given one design error. The consolidated average results from this pilot are shown in Figure 3. The horizontal axis captures the individual design errors, while the vertical axis is the average level of anticipated success as defined by the pilot students.

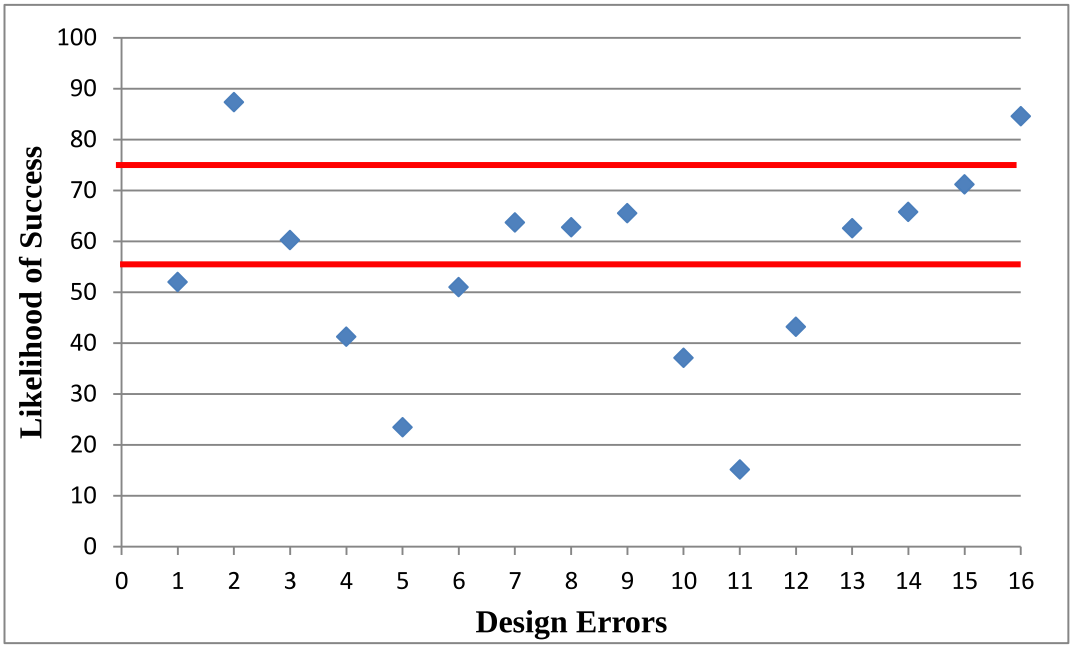


Figure 3: Assessment of Likelihood of System Success for Each Design Error

Figure 3 shows the average confidence level for the sixteen possible design errors. In order to normalize the design errors to be used for the experiment, the design errors that fit within a band of 55%-75% were selected, resulting in seven design errors. This ensures that each of the experimental design errors should be considered equivalent with relatively uniform assessment of impact on likelihood of success. The final design errors, and their associated controls, that were selected for use in the experiment are shown in Table 4.

Table 4: Design errors and controls used in the experiment

Sl #	**Design Error (and Control)**
1	Swaying of ballast while moving with an A frame to the test location Control: Re-design A frame to reduce height of the frame
2	Collection of surface material below the ballast during a drag test Control: Ensure even surface after every test
3	Irregular ground surface during drag test Control: Inspect ground surface before placing ballasts
4	Insertion of hook from the load cell to the ballast Control: Increase dimension of hole where the hook is inserted
5	Positioning of load cell during calibration Control: Check load value after every trial during calibration
6	Instability in grey frame during drag test Control: Place rubber mats below the frame to increase friction
7	Rope which prevents ballast from swaying breaks Control: Use rope of higher strength to support heavy load

3.3.4 Experimental Participants

Participants in this experimental user study were drawn from Clemson University and came from a variety of majors such as general engineering (domain general engineers), graduate mechanical engineering (domain specialists), and psychology students (non-domain individuals). These populations are illustrated in Table 5.

Table 5: Experimental Population

Population	**Background**	**Year in School**	**Number of Participants**
Domain Generalists	General Engineering	Freshman (Year 1)	117
Domain Specialists	Mechanical Engineering	Graduate (Year 5+)	23
Non- Domain Generalists	Non-Engineering	Junior (Year 3)	43

The students participated on a voluntary basis and conducted external to their scheduled class times. A design review workshop was offered to students in three different courses (general engineering, graduate mechanical engineering, and undergraduate general psychology). All participants, regardless of background, were given the same workshop presentation that included an overview of FMEA, an introduction of the basic workings of FMEA, and a detailed explanation of the tent ballast testing system (experimental problem). The training was done to assure that all participants would be capable of conducting a general design review and that they understood the challenge and criticality of assessing the severity of identified design errors. The training session was administered across seven sessions and was conducted by the same researcher in all sessions.

A design review team is typically organized based on the members' specialization and levels of expertise, rather than on general demographics. Therefore, factors relating to gender, race, socio-economic standing, and personality were not considered in selecting or organizing the participants. The variance in assessment levels due to these factors is not the focus of this experiment. The study was conducted in standard and familiar classroom settings. While the study was conducted in multiple rooms, all of the students within a given background (non-engineering, general engineering, engineering specialists) conducted the study in the same classroom.

3.3.5 Experimental Procedure

A presentation on the working of the tent testing mechanism was provided to the all the students before conducting the study. The presentation provided the participants with a basic understanding of the experimental design problem and its solution components. The participants' questions regarding the working and function of each component were answered before conducting the study. After the presentation, documents containing information about specific design errors/failure modes were provided to the individual participants. The participants were divided into two groups without regard to gender, race, or personality. The first group was provided design errors without controls and the second group was provided design errors with controls. Table 6 shows the different sets of scenarios given to the participants.

Table 6: Experiment Layout

Student Package	**Control**	**Scenario**			
		1	**2**	**3**	**4**
1	NO	A	BCA	DEBCA	FGDEBCA
2	YES	G	ABG	CDABG	EFCDABG
3	NO	B	CDB	EFCDB	GAEFCDB
4	YES	F	GAF	BCGAF	DEBCGAF
5	NO	C	DEC	FGDEC	ABFGDEC
6	YES	E	FGE	ABFGE	CDABFGE
7	NO	D	EFD	GAEFD	BCGAEFD
8	YES	D	EFD	GAEFD	BCGAEFD

9	NO	E	FGE	ABFGE	CDABFGE
10	YES	C	DEC	FGDEC	ABFGDEC
11	NO	F	GAF	BCGAF	DEBCGAF
12	YES	B	CDB	EFCDB	GAEFCDB
13	NO	G	ABG	CDABG	EFCDAGB
14	YES	A	BCA	DEBCA	FGDEBCA

Design Errors: A, B, C, D, E, F, and G

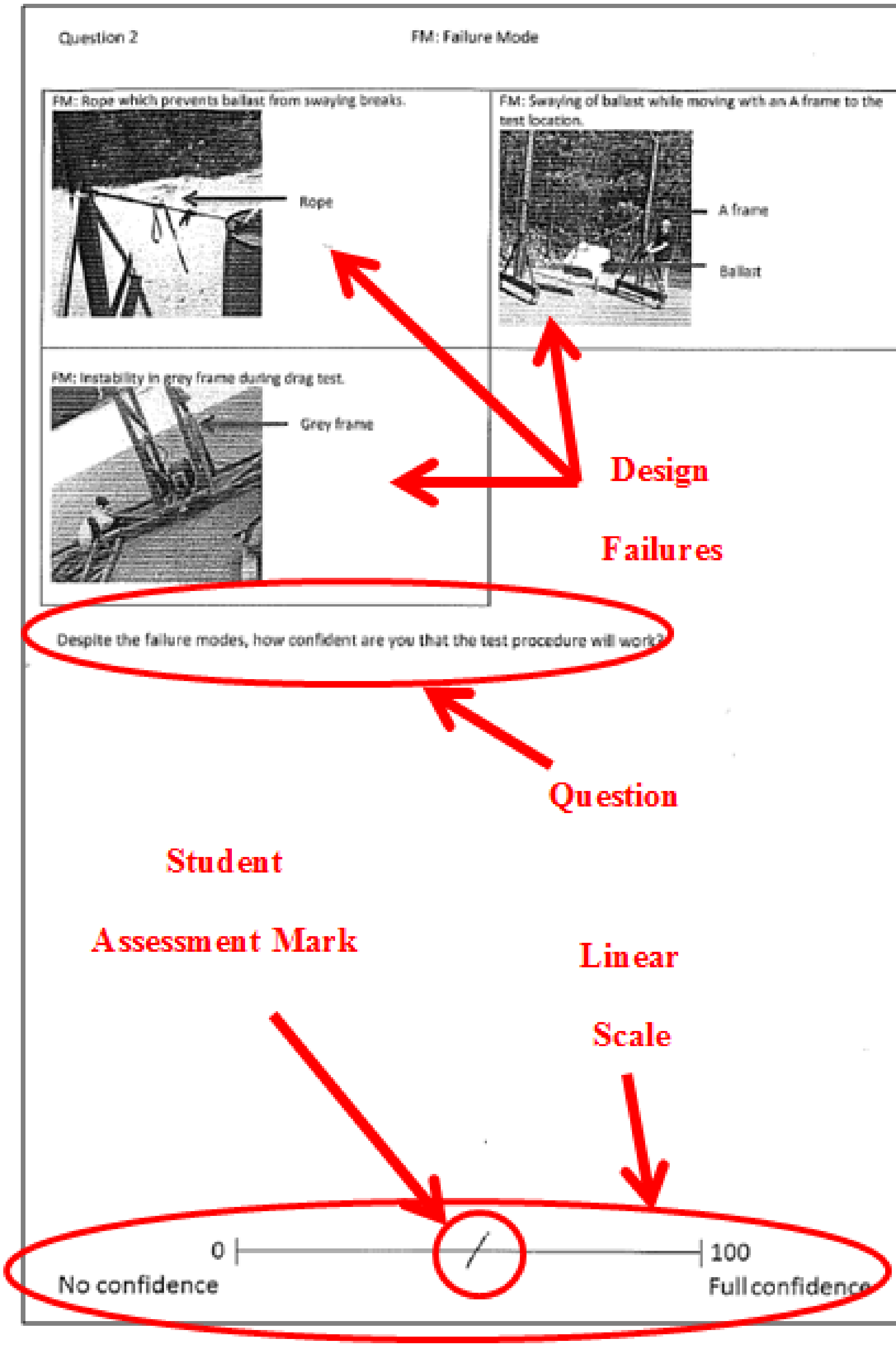
The difference in the documents is the order in which the design errors are presented. As indicated in Table 6, the design errors were presented thusly to reduce the memory effect. The design error presented in scenario 1 (*A*), is presented at the end in scenario 2 (BC*A*), the design errors presented in scenario 2 (*BCA*), are presented at the end in scenario 3 (DE*BCA*), and the design errors presented in scenario 3 (*DEBCA*), are presented at the end in scenario 4 (FG*DEBCA*). In this way each participant examines all the design errors in the experiment.

The participants, based on the number of design errors presented to them, rated their confidence level on the linear scale provided at the bottom of each scenario. The confidence rating is based on the question:

“Despite the failure mode, how confident are you that the procedure will work?”

This question remained the same for all scenarios. Each participant was given ten minutes to rate their confidence for four scenarios. The time was determined based on results from the pilot study. A sample of one of the pages from the packets given to the participants is shown in Figure 4.

Each worksheet provided the design errors for a given scenario. In order to ensure the participant fully understands the scenario, each possible error was listed individually with both an associated picture of the component in question and a textual description of the error. It should be noted that the scenario shown in Figure 4 does not include controls for the given design errors. The confidence “slider bar” was provided at the bottom of the worksheet and the student participants mark their confidence on the scale.



Question 2

FM: Failure Mode

FM: Rope which prevents ballast from swaying breaks.

FM: Swaying of ballast while moving with an A frame to the test location.

FM: Instability in grey frame during drag test.

Despite the failure modes, how confident are you that the test procedure will work?

0 No confidence

100 Full confidence

Figure 4: Example of Error/ Failure Mode Worksheet (Scenario 2)

3.3.6 Data Collection

The data from the study consists of the confidence ratings of each participant for the four scenarios. Each document was evaluated individually by measuring on a continuous scale. A linear scale, similar to the 100-mm line discussed above, was used as the measuring scale in this study. This scale was chosen because the use of the 100-mm line has been proven to be advantageous when conducting experiments with a large number of participants. The linear scale is also useful for obtaining precise measurements as there is no preset intervals towards which the participants would tend, such as marks at increments of 5%, 10%, or 25%. As a result, the participants are more likely to provide precise responses without any rounding bias.

For this measure, participants answered the question, “Despite the failure mode, how confident are you that the procedure will work?” The participants were instructed to indicate their confidence by making a slash mark on the linear scale indicating their confidence from no confidence to full confidence. Each participant’s response was measured on a continuous scale and recorded using a standard ruler. The distance is measured from the left start point of the linear scale to the point where the slash mark intersects the linear scale, as shown in the bottom of Figure 4. The distance is then converted to a percentage of the full scale that is used for analysis. The participants’ confidence ratings for each scenario are then combined and averaged in order to determine the confidence level for project success considering various numbers of design errors both with and without controls.

CHAPTER FOUR: ANALYSIS AND RESULTS

The average of the confidence ratings for multiple failure modes are compared for three domain populations. Significant statistics are derived to indicate the influence of individuals' domain, and the influence of providing controls (solutions) to the errors in estimating system success. By analyzing the documents and comparing the confidence ratings (5.4.6Appendix B:), the influence of failure number, specialization and domain on confidence in estimating system performance during design review is studied.

4.1 Results on Domain and Design Error Interaction

Figure 5 indicates the decrease in confidence for the three domain populations as the number of design errors increase. The confidence level of the domain population indicated in Figure 5 is a combination of design errors with and without controls. In the beginning, the confidence level appears to be the same. However, as the number of design errors increase, we can notice a downward trend for all three domain populations, and the variation of confidence level between the specialists and generalists.

For a single design error (Table 7), there is insufficient evidence to conclude that there is a difference in the mean confidence between the three domain populations (p=0.3456), indicating that the confidence level is the same, or at least the means are not significantly different. The three domain populations start at the same level, which suggests that the seven design errors considered for the experiment are comparable, as discussed in Section 3.3.3. At this point it does not matter who the decision maker is since the confidence levels in assessing the performance of the system for all domain population are comparable.

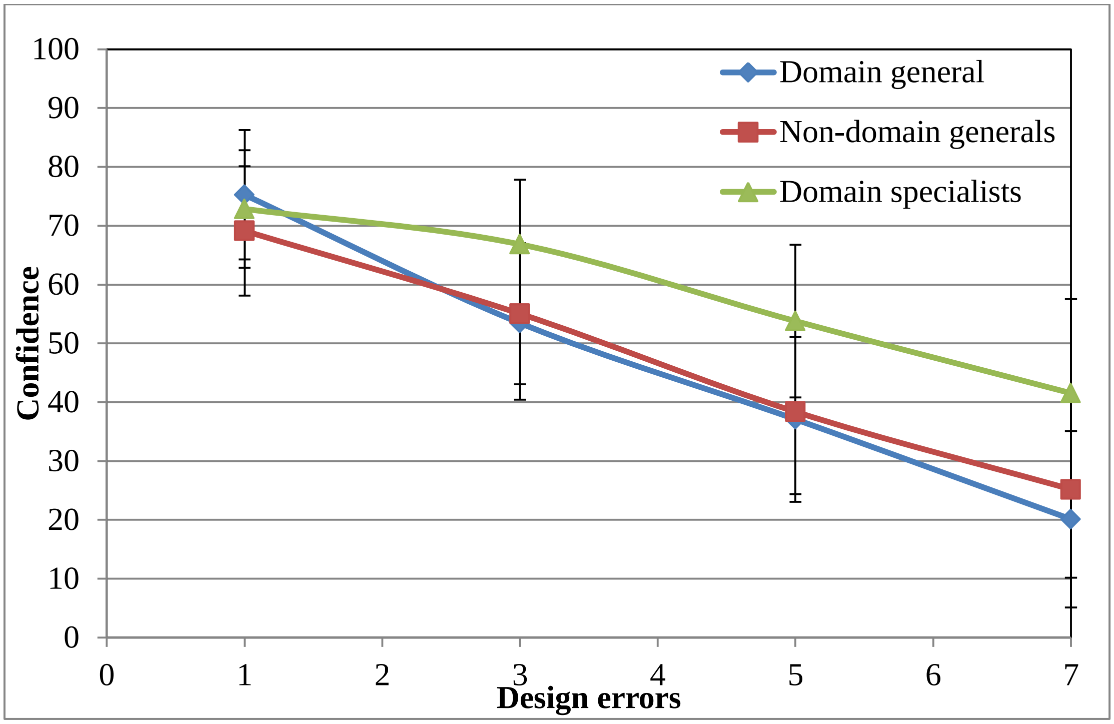


Figure 5: Confidence level comparison between different domain populations

For three design errors (Table 7) there is evidence of a difference in the mean confidence among the three domain populations at a level of significance 10% (p=0.0793). The mean confidence for the domain specialists is higher than domain generalists and non-domain generalists for three design errors. Beyond single design error, a downward trend is observed for all domain population and the difference in the confidence level is evident. When the number of design errors is three or more, care should be taken during decision making since the difference in the confidence level for assessing the system performance among the three domain population is evident.

Specialists showed optimism in predicting the performance of the system for three design errors.

For five design errors (Table 7) at a 5% level of significance there is evidence of a difference in the mean confidence among the three domain types (p=0.0198). The mean confidence for the domain specialists is higher than domain generalists and non-domain generalists. As indicated by the p-value, the difference in the confidence level is more evident as compared to the case of three design errors. This suggests that, as the number of failure modes considered at a given time increases, the difference in confidence level during decision making is larger.

For seven design errors (Table 7) at a 5% level of significance there is evidence of a difference in the mean confidence among the three domain types (p=0.0015). The mean confidence for the domain specialists is higher than domain generalists and non-domain generalists. The trend in confidence level from a single design error to seven design errors suggests that the domain specialists are more optimistic in predicting the performance of the system.

Results in Table 7 suggest that *when considering one design error at a time, it does not matter if the decision maker is a domain generalist or a domain specialist, they can be considered equivalent. This condition is similar to the investigation of the failure modes conducted during FMEA, considering only single errors at a time. However, while considering more than one design error at a given time (multiple failure modes), there is a significant difference in the decision (Table 7). Thus, care should be given to decisions when considering multiple errors simultaneously.*

Table 7: Effect of design errors on all domain population

# of design errors	Domain generalists	Std Dev.	Non-domain generalists	Std Dev.	Domain specialists	Std Dev.	p-value
1	75.26	11	69.13	11	72.86	10	0.3456
3	53.45	13	55.04	12	66.84	11	0.0793
5	37.09	14	38.38	14	53.79	13	0.0198
7	20.12	15	25.19	15	41.54	16	0.0002

Table 8 indicates the rate of decrease in confidence for the three domain population. The confidence level indicated for multiple design errors is a combination of design errors with and without their associated controls. The decrease in confidence is faster for domain generalists with a p-value less than 0.0001. The rate of decrease in confidence is the least for the domain specialists, who appear to show a higher confidence level as design errors increase when compared to the generalists. The plummet in the confidence level of the non-domain generalists and domain generalists indicate they are pessimistic about the performance of the system while considering multiple failure modes at the same time. However, the domain specialists remain optimistic about the performance of the system. The specialists appear to have confidence in the success of system performance or at least estimates a positive outcome when multiple failure modes are considered. This suggests that, *if the specialist gives a negative estimation of the system success, they can be trusted. In other words, the domain generalists can be trusted when they give a positive estimation of the system success and cannot be trusted when they give a negative response.*

Table 8: Percent Decrease in Confidence

Domain	Design Error	% decrease in confidence	p-value
Specialists	1-3	6.1	0.2802
Specialists	3-5	13	0.0211
Specialists	5-7	12.2	0.0301
Non-domain	1-3	14.1	0.0002
Non-domain	3-5	16.7	<0.0001
Non-domain	5-7	13.2	0.0005
General	1-3	21.8	<0.0001
General	3-5	16.4	<0.0001
General	5-7	16.9	<0.0001

4.1.1 Domain Specialists

The confidence level in assessing the performance of the system did not decrease significantly from a single design error to three design errors. However, confidence level decreased significantly beyond three design errors. The estimated decrease in mean confidence from a single design error to three design errors is 6% (Table 8), however there is insufficient evidence to conclude that this decrease is significant (p=0.2802). *The specialists trusted the performance of the system even though three design errors were considered at the same time indicating optimism in predicting the performance.*

The estimated decrease in mean confidence from three design errors to five design errors and five design errors to seven design errors is 13% (p=0.0211) & 12.2% (p=0.0301) respectively. There is sufficient evidence at 5% level of significance to conclude the decrease is significant (Table 8). The decrease in confidence is evident when considering more than three design errors. This is when the specialists' trust in predicting the performance of the system reduces. A downward trend was observed as more failure modes were considered. However, this downward trend is at a higher confidence level compared to the generalists.

4.1.2 Non-Domain Generalists

The confidence level for non-domain generalists decreased significantly for multiple design errors (Table 8). The estimated decrease in mean confidence from a single design error to three design errors, three design errors to five design errors, and five design errors to seven design errors is 14.1% (p=0.0002), 16.7% (p<0.0001), and 13.2% (p=0.0005) respectively. There is sufficient evidence at 5% level of significance to conclude this decrease in mean confidence level for multiple design errors is significant (Table 8). *The non-domain generalists' trust in predicting the performance significantly reduced for multiple design errors*.

4.1.3 Domain Generalists

The confidence level decreased significantly for multiple design errors (Table 8). The decrease in confidence level is faster for domain generalists compared to non-domain generalists and domain specialists. The estimated decrease in mean confidence level from a single design error to three design errors, three design errors to five design errors, and five design errors to seven design errors is 21.8% (p<0.0001), 16.4% (p<0.0001), and 16.9% respectively. There is sufficient evidence at a 5% level of significance to conclude the decrease is significant.

The decrease in the confidence level for domain generalists from design errors 1-3, 3-5, and 5-7 indicate they were the most pessimistic in predicting the system performance ($p<0.0001$). This suggests that *the domain generalists cannot be trusted with predicting the system performance if they give a negative opinion. Their estimation on the success of the system can be trusted only if they indicate a positive outcome*.

4.2 Results for Increase in Confidence Using Controls

The influence of providing controls or solutions to the design errors, on the confidence level is examined for multiple design errors considered at the same time. Figure 6 illustrates the impact controls have in increasing the overall confidence of the participants regardless of the number of design errors introduced. The confidence level indicated in Figure 6 is a combination of all domain populations. For a single design error the increase in confidence is not evident. However, while considering multiple design errors, controls proved significant in increasing the confidence level in predicting the system performance.

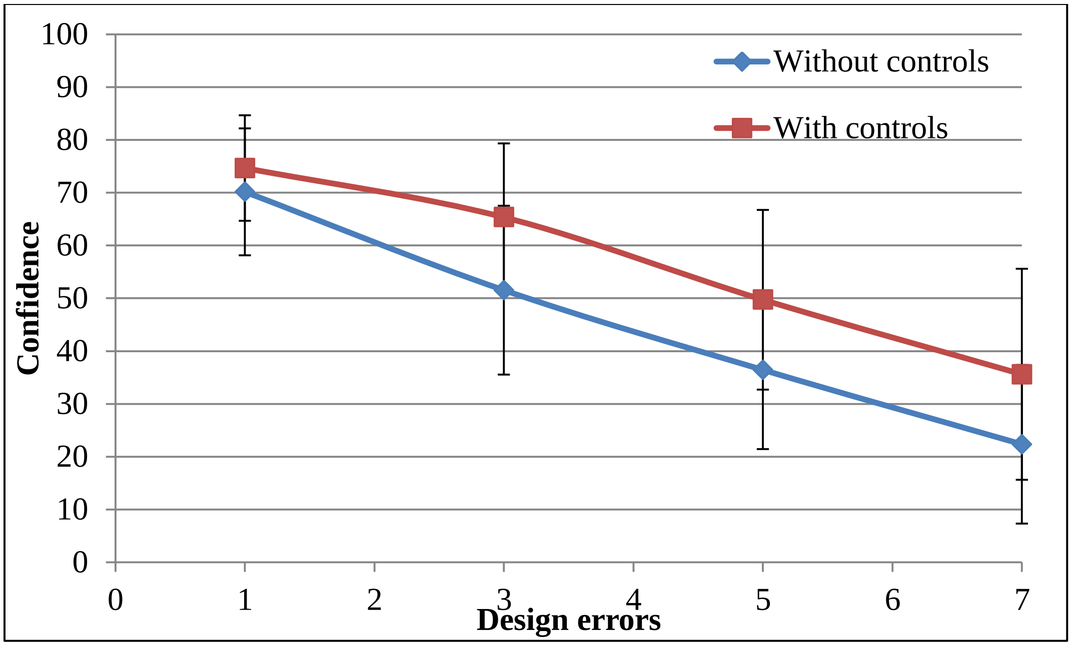


Figure 6: Effect of Controls on all Domain population

For a single design error the increase in confidence with control is 4.5%. It is estimated that there is insufficient evidence to conclude there is significant increase in confidence using controls for a single design error (p=0.3348). The confidence level did not increase significantly by using control for a single design error. This condition is similar to the investigation of the failure modes conducted during FMEA, considering only single errors at a time. *Providing control for a single error does not significantly improve the confidence level in predicting the system performance.*

For three design errors, five design errors, and seven design errors the increase in confidence level is 13.8% (p=0.0034), 13.3% (p=0.0048), and 13.2% (p=0.0049) respectively. It is estimated that at a 5% level of significance, there is sufficient evidence to conclude there is significant increase in confidence using controls for multiple design errors. Providing controls for multiple errors increased the confidence level significantly in predicting the system performance.

The increase in confidence level due to the use of control was evident only while considering multiple design errors at the same time. This condition compared to the investigation of the failure modes conducted during FMEA, providing controls to a single design error does not significantly increase the confidence in predicting the system performance. *Providing controls to multiple failure modes increased the confidence level significantly. However, the confidence level with the use of control had a downward trend.*

Table 9: Effect of Controls on all Domain Population

Design Error	% confidence without control	Std Dev.	% confidence with control	Std Dev.	p-value
1	70.16	12	74.67	10	0.3348
3	51.54	16	65.35	14	0.0034
5	36.44	15	49.73	17	0.0048
7	22.35	15	35.61	20	0.0049

The confidence level for all domain population at any number of design errors increase when the associated controls are provided to the design errors. Table 10 indicates the confidence level in predicting system performance for all domain population with and without providing the associated controls for the design errors.

Table 10: Confidence level comparison for all domain population with and without controls

Error	Population	Average Percent confidence without control	Std Dev.	Average Percent confidence with control	Std Dev.	P-value
1	Domain Generalists	74	11	76	12	0.6794
3		50	13	57	13	0.0926
5		32	13	42	15	0.0089
7		15	14	25	15	0.0134
1	Non-Domain Generalists	68	13	70	9	0.7558
3		45	13	65	12	0.0066
5		33	15	43	14	0.1520
7		21	16	29	15	0.3201
1	Domain Specialist	66	12	76	7	0.3916
3		57	12	72	8	0.2053
5		42	13	62	10	0.0944
7		28	16	51	17	0.0448

4.2.1 Domain Generalists

Figure **7** illustrates the impact of introducing controls on the domain generalist population. For a single design error (Figure 7), at a 10% level of significance, there is insufficient evidence to conclude there is a difference in the mean confidence using controls (p=0.6794). The confidence level in predicting the system performance using control for a single design error is considered comparable. It does not matter if control is being used for a single design error since the confidence level in assessing the performance of the system is comparable.

For three design errors (Figure 7), at a 10% level of significance, there is sufficient evidence to conclude there is a difference in the mean confidence using controls (p=0.0926). The confidence level in predicting the system performance using control for three design errors is higher than the confidence level without using control. It does matter if control is being used for three design errors since the confidence level in assessing the performance of the system is higher when control is introduced.

For five design errors and seven design errors (Figure 7), at a 5% level of significance, there is sufficient evidence to conclude there is a difference in the mean confidence level when using controls. The confidence level in predicting the system performance using the associated control for five design errors and seven design errors is higher than the confidence level without using the control. It does matter if control is being used for five design errors and seven design errors since the confidence level in assessing the performance of the system is higher when control is introduced.

For multiple design errors, the introduction of controls increased the confidence level for the domain generalists in estimating the system performance. However, the introduction of controls did not reduce the rate of decrease in confidence. This suggests that the use of controls did not affect the sense of pessimism in estimating the system success. The reason for this is the lack of engineering expertise in understanding the implications of using controls.

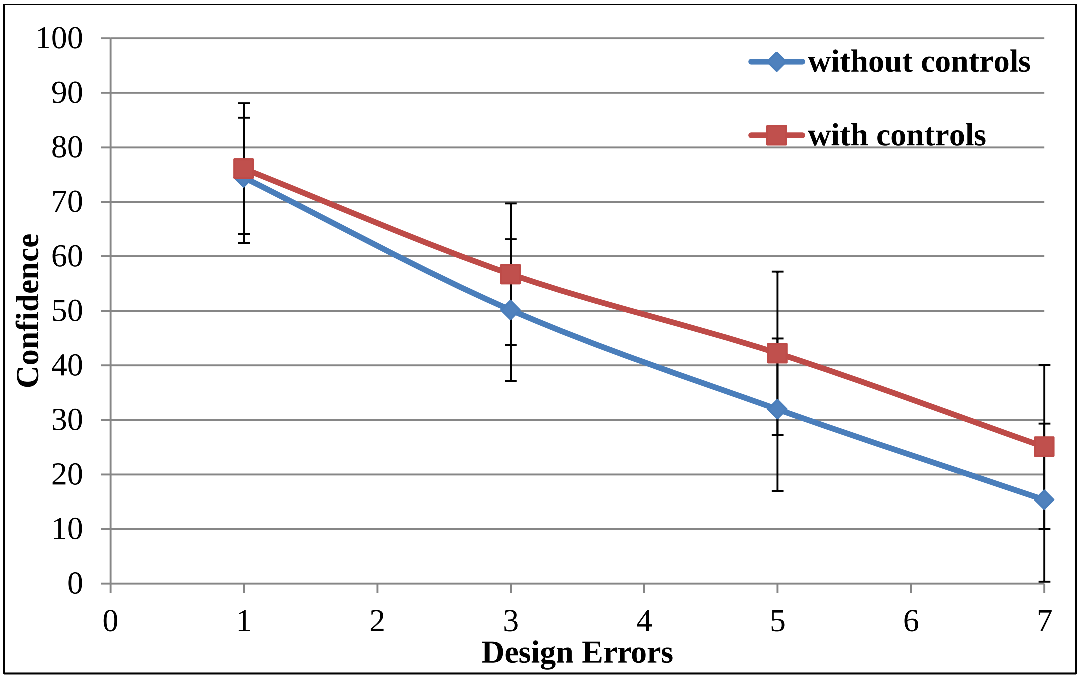


Figure 7: Confidence vs. Design errors for Domain Generalists with and without Controls

4.2.2 Non-Domain Generalists

Figure 8 illustrates the impact of controls on the non-domain generalist population of psychology class students. For a single design error (Figure 7), at a 10% level of significance, there is insufficient evidence to conclude there is a difference in the mean confidence using controls (p=0.7558). The confidence level in predicting the system performance using control for a single design error is considered comparable. It does not matter if control is being used for a single design error since the confidence level in assessing the performance of the system is comparable.

For three design errors (Figure 7), at a 5% level of significance, there is sufficient evidence to conclude there is a difference in the mean confidence using controls (p=0.0066). The confidence level in predicting the system performance using control for three design errors is higher than without control. It does matter if control is being used for three design errors since the confidence level in assessing the performance of the system is higher when the associated control is introduced.

For five design errors and seven design errors (Figure 7), at a 10% level of significance, there is insufficient evidence to conclude there is a difference in the mean confidence level using controls (p=0.1520 & 0.3201 respectively). The confidence level in predicting the system performance using control for five design errors and seven design errors is comparable with the confidence level without using control.

The introduction of controls increased the confidence level in estimating the system performance for the non-domain generalists. The confidence level for design errors one and three is comparable, beyond this the confidence level decreased drastically, indicating a point of transition. Due to the rate of decrease in confidence, we can conclude that the non-domain generalists are not optimistic about the system success. The lack of technical knowledge could be a reason for the above mentioned findings. *The trend in the confidence level for both, domain generalists and non-domain generalists can be considered comparable since the domain generalists have only 5-6 months of engineering knowledge and the non-domain generalists have almost no engineering knowledge over their course of study.*

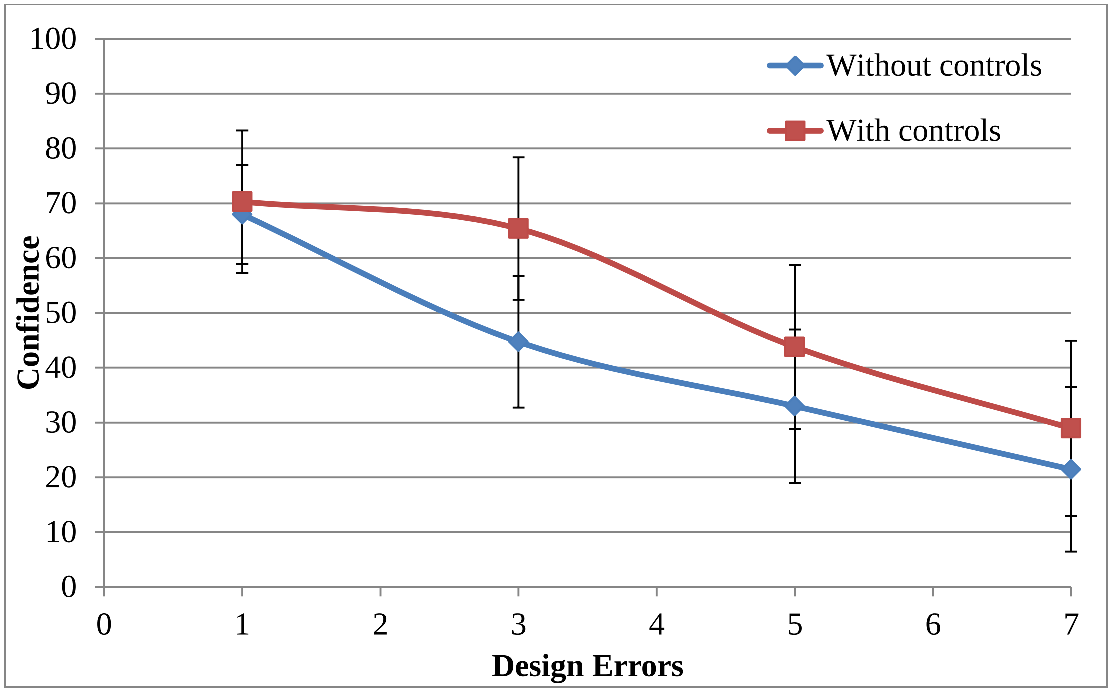


Figure 8: Confidence vs. design errors for Non-Domain generalists with and without Controls

4.2.3 Domain Specialists

Figure 9 illustrates the impact of controls on the domain specialist population. For a single design error and three design errors (Figure 7), at a 10% level of significance, there is insufficient evidence to conclude there is a difference in the mean confidence using controls (p=0.3916 & 0.2053 respectively). The confidence level in predicting the system performance using control for a single design error and three design errors is considered comparable. It does not matter if control is being used for a single design error or three design errors since the confidence level in assessing the performance of the system is comparable.

For five design errors and seven design errors (Figure 7), at a 10% level of significance, there is sufficient evidence to conclude there is a difference in the mean confidence using controls (p=0.0944 & 0.0448 respectively). The confidence level in predicting the system performance using control for five design errors and seven design errors is higher than without the use of control. It does matter if control is being used for five design errors or seven design errors since the confidence level in assessing the performance of the system is higher when control is introduced.

It can be seen that there is an increase in confidence for all populations when controls are introduced for multiple failure modes. The shape of the trends suggests that the domain generalists and specialists have near linear relationships with the number of increasing design errors. The non-domain generalists do not show such a linear trend. The transition point or inflection point can be observed for non-domain generalists beyond three design errors. The same could not be observed for the domain generalists and domain specialists. It can be seen that the decrease in confidence level for domain generalists is sharp and hence, the confidence level could be very low before we can observe the transition point. However, due to the trend observed for the domain specialists suggests that the point of transition could be observed by introducing more design errors, before the confidence level drops significantly. *The knowledge or domain awareness of the specialists enables them to make an optimistic estimation of the system performance for multiple errors.*

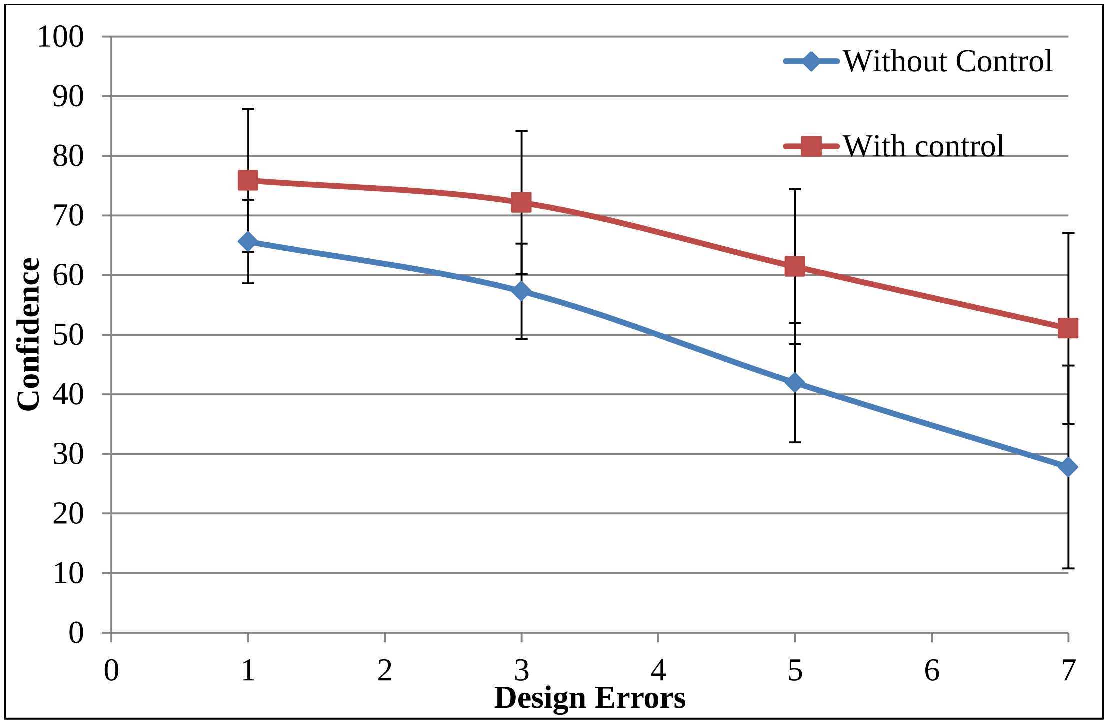


Figure 9: Confidence vs. design errors for Domain Specialists with and without Controls

4.3 <u>Pairwise comparisons</u>

Pairwise comparisons are made between the domain generalists, non-domain generalists and domain specialists to judge which of the population is preferred, or has a greater influence in decision-making.

4.3.1 Domain Generalists vs. Non-Domain Generalists

A comparison is made between the domain generalists and non-domain generalists to study the difference between individuals with minimum engineering background and individuals without any engineering background. The confidence level indicated in Figure 10 is a combination of design errors with and without controls. Figure 10 indicates the downward trend in the confidence level for the domain generalists and non-domain generalists as the design error increases. The confidence level for both the populations indicated is a combination of design errors with and without controls. The confidence level appears to be the same for multiple design errors.

4.3.1.1 Domain Generalists vs. Non-Domain Generalists – Combined (with and without controls)

For a single design error, and multiple design errors (Figure 10), there is insufficient evidence to conclude there is a difference in the mean confidence among domain generalists and non-domain generalists (Table 11), indicating that the confidence level is the same. During assessment of the system performance for a single design error or multiple design errors, a domain general and a non-domain general are considered comparable.

Results in Figure 10 indicate *there is no significant difference between the two domain populations. This result is validated by the fact that the domain generalists have an engineering experience of less than six months. Hence the domain generalists and the non-domain generalists can be considered to be pessimists in predicting the system success. Their decision can be trusted if they provide a positive outcome on the performance of the system.*

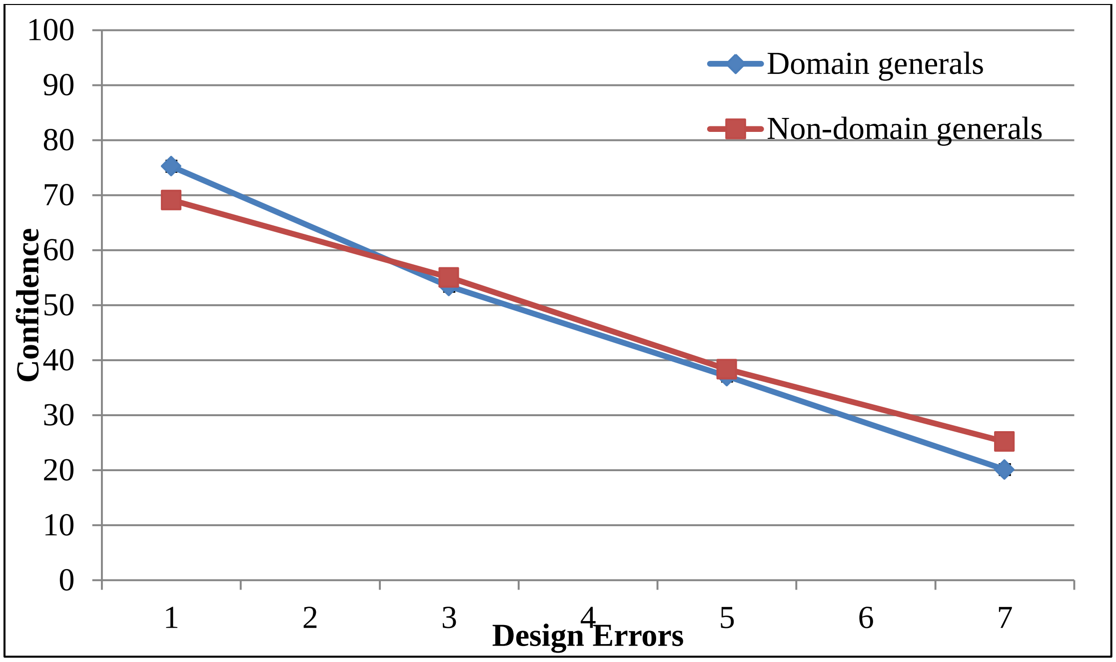


Figure 10: Domain Generalists vs. Non-Domain Generalists - Combined (with & without controls)

Table 11: Confidence level for Domain Generalists vs. Non-Domain Generalists - Combined (with & without controls)

# of Design Errors	Domain Generalists	Non-Domain Generalists	P-Value
1	75.26	69.13	0.1488
3	53.45	55.04	0.7064
5	37.09	38.38	0.7615
7	20.12	25.19	0.2393

4.3.1.2 Domain Generalists vs. Non-Domain Generalists – With controls

A comparison is made between the domain generalists and non-domain generalists to estimate the system success for design errors with control. Figure 11 indicates the downward trend in the confidence level for the domain generalists and non-domain generalists as the design error increases. The confidence level for both the populations indicated is for design errors with controls. The confidence level appears to be the same for multiple design errors.

For a single design error and multiple design errors (Figure 11), there is insufficient evidence to conclude there is a difference in the mean confidence among domain generalists and non-domain generalists (Table 12), indicating that the confidence level is the same. During assessment of the system performance for a single design error or multiple design errors with control, a domain general and a non-domain general are considered comparable.

Results in Table 12 indicate that the use of controls for the design errors did not differentiate the domain generalists and non-domain generalists. The reason behind minimum increase in confidence for both domain populations with the use of controls is the lack of knowledge to understand the implication of using controls. A larger difference can be seen for three design errors since the non-domain generalists have a transition point beyond three design errors.

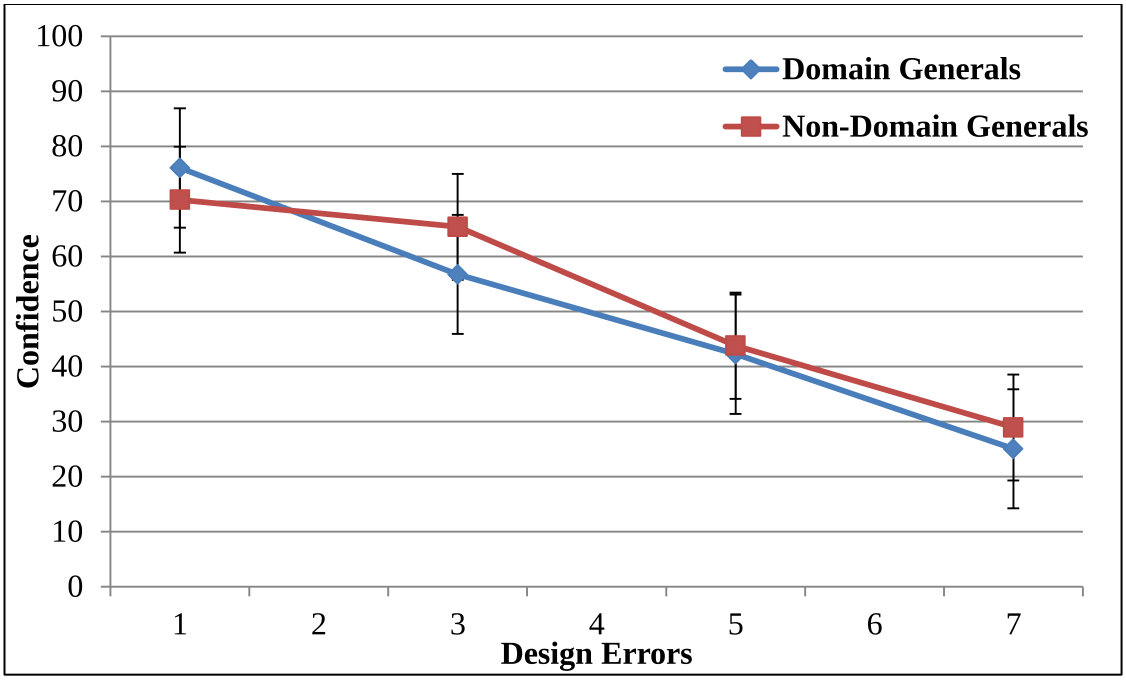


Figure 11: Domain Generalists vs. Non-Domain Generalists – With controls

Table 12: Confidence level for Domain Generalists vs. Non-Domain Generalists - With controls

# of Design Errors	Domain Generalists	Non-Domain Generalists	P-Value
1	76.06	70.3	0.3405
3	56.73	65.37	0.1536
5	42.23	43.78	0.7975
7	25.05	28.93	0.5203

4.3.1.3 Domain Generalists vs. Non-Domain Generalists – Without controls

A comparison is made between the domain generalists and non-domain generalists in estimating the system performance for design errors without the use of controls. Figure 12 indicates the downward trend in the confidence level for the domain generalists and non-domain generalists as the design error increases. The confidence level for both the populations indicated is for design errors without controls. The confidence level appears to be the same for multiple design errors.

For a single design error, and multiple design errors (Figure 12), there is insufficient evidence to conclude there is a difference in the mean confidence among domain generalists and non-domain generalists (Table 13), indicating that the confidence level is the same. During assessment of the system performance for a single design error or multiple design errors without control, a domain general and a non-domain general are considered comparable.

Results suggest that when considering any number of design errors (with controls, without controls, or combined), it does not matter if the decision maker is a domain general or a non-domain general, they are considered equivalent. An individual with minimum exposure to engineering (domain generalists) is equivalent to an individual with no engineering experience (non-domain generalists). A positive outcome from the two domain populations in predicting the performance of a system can be trusted, in other words, a domain specialist with negative outcome on the performance of a system cannot be trusted.

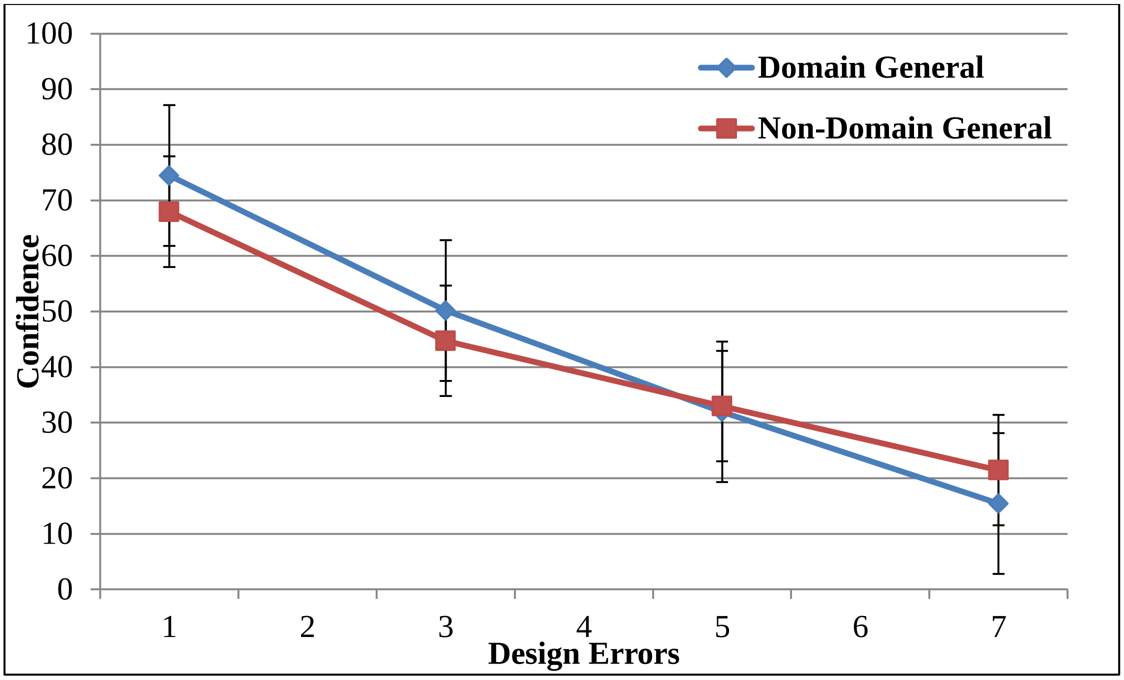


Figure 12: Domain Generalists vs. Non-Domain Generalists – Without controls

Table 13: Confidence level for Domain Generalists vs. Non-Domain Generalists - Without controls

# of Design Errors	Domain Generalists	Non-Domain Generalists	P-Value
1	74.46	67.96	0.2746
3	50.16	44.71	0.3592
5	31.95	32.97	0.8635
7	15.45	21.45	0.3049

4.3.2 Domain Generalists vs. Domain Specialists

A comparison is made between the domain generalists and domain specialists to study the difference in estimating system performance for design errors with controls, without controls and combination of design errors with and without design errors. Figure 13 indicates the downward trend in the confidence level for the domain generalists and domain specialists as the design error increases. The confidence level for both the populations indicated is a combination of design errors with and without controls.

4.3.2.1 Domain Generalists vs. Domain Specialists – Combined (with & without controls)

For a single design error (Table 14), there is insufficient evidence to conclude there is a difference in the mean confidence among domain generalists and domain specialists (0.6856), indicating that the confidence level is the same. During assessment of the system performance for a single design error, a domain general and a domain specialist are considered comparable.

For three design errors, five design errors and seven design errors (Figure 13), at a 5% level of significance there is sufficient evidence to conclude there is a difference in the mean confidence among domain generalists and domain specialists (Table 14), indicating that the confidence level is not the same. During assessment of the system performance for multiple design errors, care should be taken during decision making between a domain general and a domain specialist.

Results in Table 14 indicate the difference in estimating system success between the domain generalists and domain specialists as the number of design errors increase. The difference in estimation is larger as indicated by the p-values as the errors increase.

The specialists are optimistic about the performance of the system, hence, their decision for multiple failure modes can be trusted if it has a negative outcome. The domain specialists are pessimistic about the system performance and their decision can be trusted if it has a positive outcome.

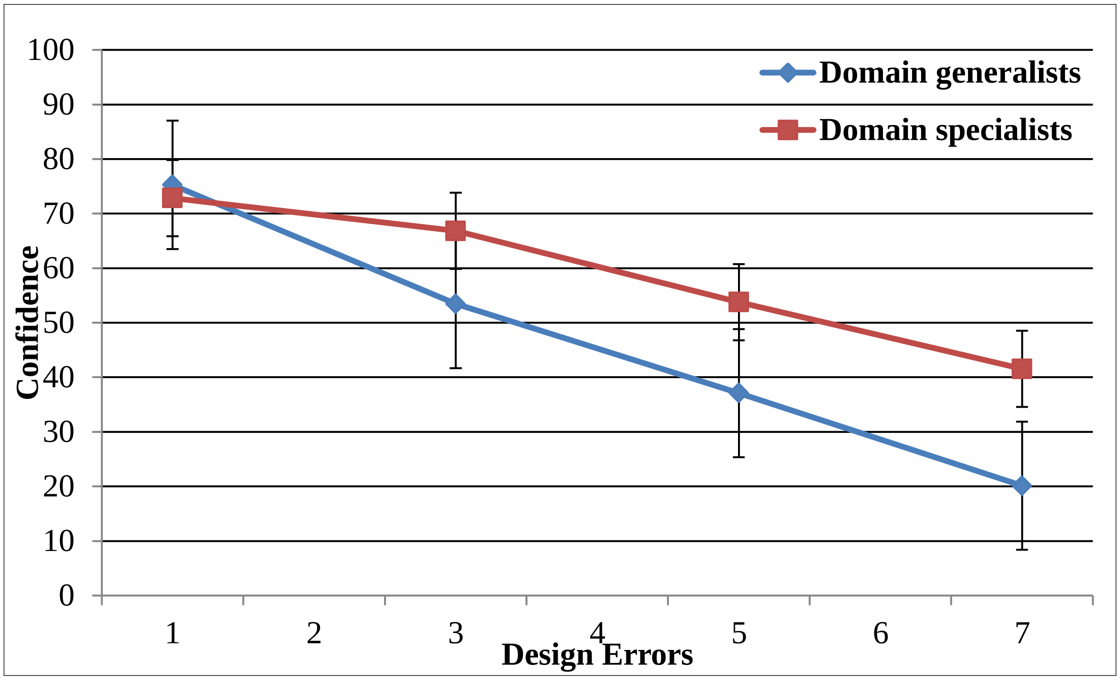


Figure 13: Domain Generalists vs. Domain Specialists - Combined (with & without controls)

Table 14: Confidence level for Domain Generalists vs. Domain Specialists - Combined (with & without controls)

# of Design Errors	Domain Generalists	Domain Specialists	P-Value
1	75.26	72.86	0.6856
3	53.45	66.84	0.0247
5	37.09	53.76	0.0053
7	20.12	41.54	0.0004

4.3.2.2 Domain Generalists vs. Domain Specialists – With controls

For a single design error (Table 15), there is insufficient evidence to conclude there is a difference in the mean confidence among domain generalists and domain specialists (0.8501), indicating that the confidence level is the same. During assessment of the system performance for a single design error with control, a domain generalist and a domain specialist are considered comparable.

For multiple design errors (Figure 14), at a 5% level of significance there is sufficient evidence to conclude there is a difference in the mean confidence among domain generalists and domain specialists (Table 15), indicating that the confidence level is not the same. The difference in confidence level between the two population types increase as more number of design errors are considered. During assessment of the system performance for three design errors with controls, care should be taken during decision making between a domain generalist and a domain specialist.

Results suggest that when considering multiple design errors with controls, care should be taken during decision making as it does matter if the decision maker is a domain generalist or a domain specialist, they are not considered equivalent. Higher confidence level for the domain specialists suggest that the domain specialists are more optimistic about the system performance and they can better understand the implications of using controls. However, there is a possibility that beyond seven design errors, the confidence level in estimation of system performance could have a point of inclination, beyond which the confidence level decreases drastically.

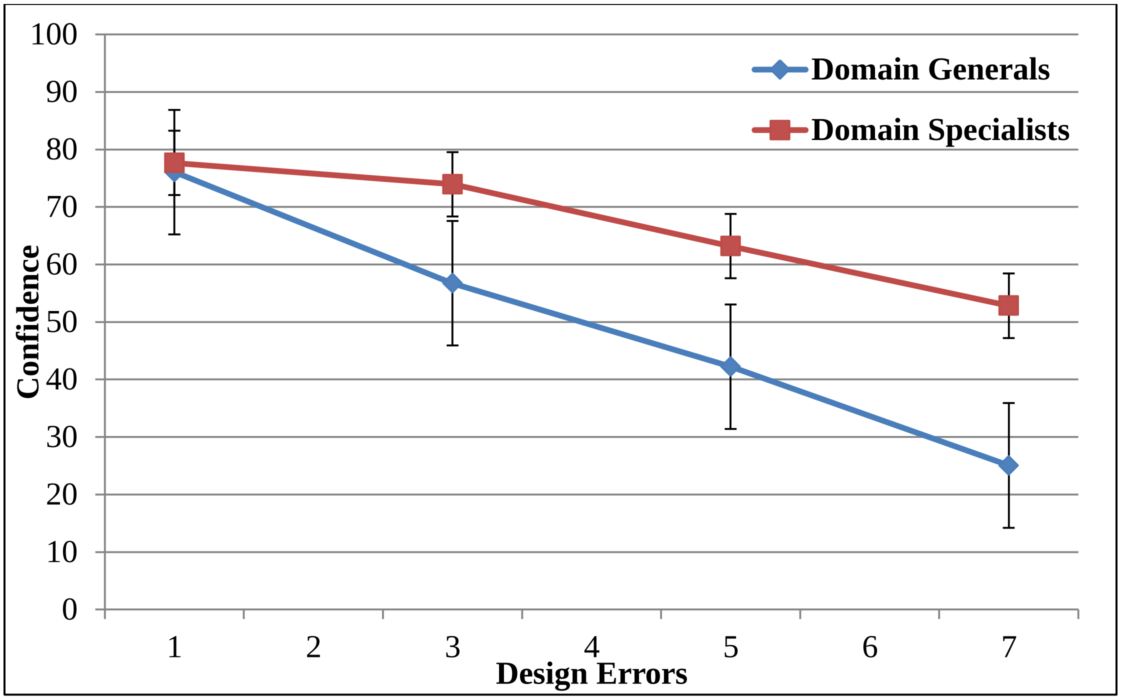


Figure 14: Domain Generalists vs. Domain Specialists – With controls

Table 15: Confidence level for 4.3.2.2 Domain Generalists vs. Domain Specialists – With controls

# of Design Errors	Domain Generalists	Domain Specialists	P-Value
1	76.06	77.66	0.8501
3	56.73	73.94	0.0426
5	42.23	63.19	0.0138
7	25.05	52.83	0.0012

4.3.2.3 Domain Generalists vs. Domain Specialists – Without controls

Figure 15 indicates the downward trend in the confidence level for the domain generalists and domain specialists as the design error increases. The confidence level for both the populations indicated is for design errors without controls.

For a single design error, three design errors and five design errors (Figure 15), there is insufficient evidence to conclude there is a difference in the mean confidence among domain generalists and domain specialists (Table 16), indicating that the confidence level is the same. For seven design errors, at a 10% level of significance there is sufficient evidence to conclude there is a difference in the mean confidence among domain generalists and domain specialists (0.0740), indicating that the confidence level is not the same.

Results suggest that when considering multiple design errors with controls, care should be taken during decision making as it does matter if the decision maker is a domain generalist or a domain specialist, they are not considered equivalent. The specialists were optimist in predicting the system performance when controls were provided to the errors. Hence a greater difference could be identified when controls were introduced. *While decision-making in terms of predicting the performance for design errors with controls, the specialists' decision can be trusted when a negative outcome is expressed since they have a better understanding about the implications of the control. The domain generalists' estimation for design errors with controls did not increase their level of confidence and they are still considered pessimists. Their decision can be trusted when they give a positive outcome for the system success.*

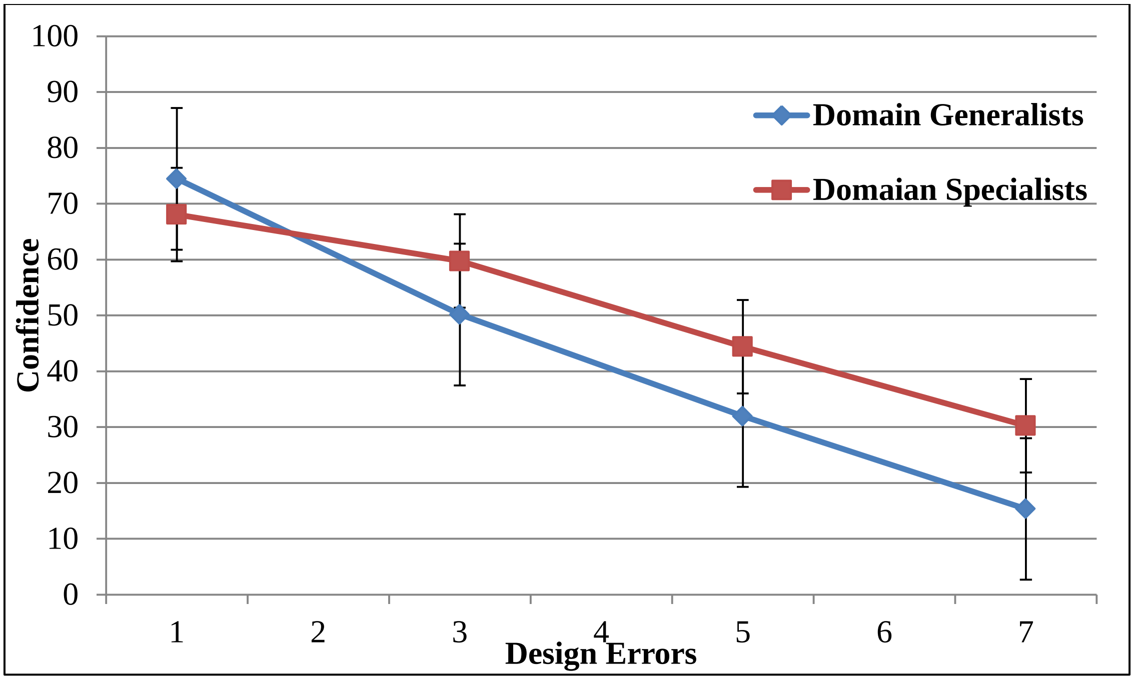


Figure 15: Domain Generalists vs. Domain Specialists - Without controls

Table 16: Confidence level for Domain Generalists vs. Domain Specialists - Without Controls

# of Design Errors	Domain Generalists	Domain Specialists	P-Value
1	74.46	68.07	0.4423
3	50.16	59.74	0.2498
5	31.95	44.4	0.1353
7	15.35	30.26	0.0740

4.3.3 Experience (Domain Specialists) vs. Minimum Experience (Domain Generalists & Non-Domain Generalists)

A comparison is made between the domain specialists against the domain generalists and non-domain generalists combined, to determine if the engineering experience of the individual makes a difference while estimating the system performance.

4.3.3.1 Experience (Domain Specialists) vs. Minimum Experience (Domain Generalists & Non-Domain Generalists) – Combined (with and without controls)

Figure 16 indicates the downward trend in the confidence level for the generalists and specialists as the design error increases. The confidence level for both the populations indicated is a combination of design errors with and without controls. For a single design error (Table 17), there is insufficient evidence to conclude there is a difference in the mean confidence among the generalists and the specialists (0.9111), indicating that the confidence level is the same. During assessment of the system performance for a single design error, generalists and specialists are considered comparable.

For multiple design errors (Table 17Figure 16), at a 5% level of significance there is sufficient evidence to conclude there is a difference in the mean confidence among the generalists and the specialists (Table 17), indicating that the confidence level is not the same. During assessment of the system performance for multiple design errors, care should be taken during decision making between an individual with experience and an individual with minimum experience.

Results suggest that when considering multiple design errors, care should be taken during decision making as the experience of the individual does matter. However, for a single design error there was no significant difference in the estimation. Specialists (maximum experience) were more optimistic about the performance of the system when multiple design errors were considered. Hence *the decision made by an individual with maximum experience on the system performance for multiple design errors can be trusted if the estimation has a negative outcome. In case of a positive outcome, the domain generalists and non-domain generalists can be trusted.*

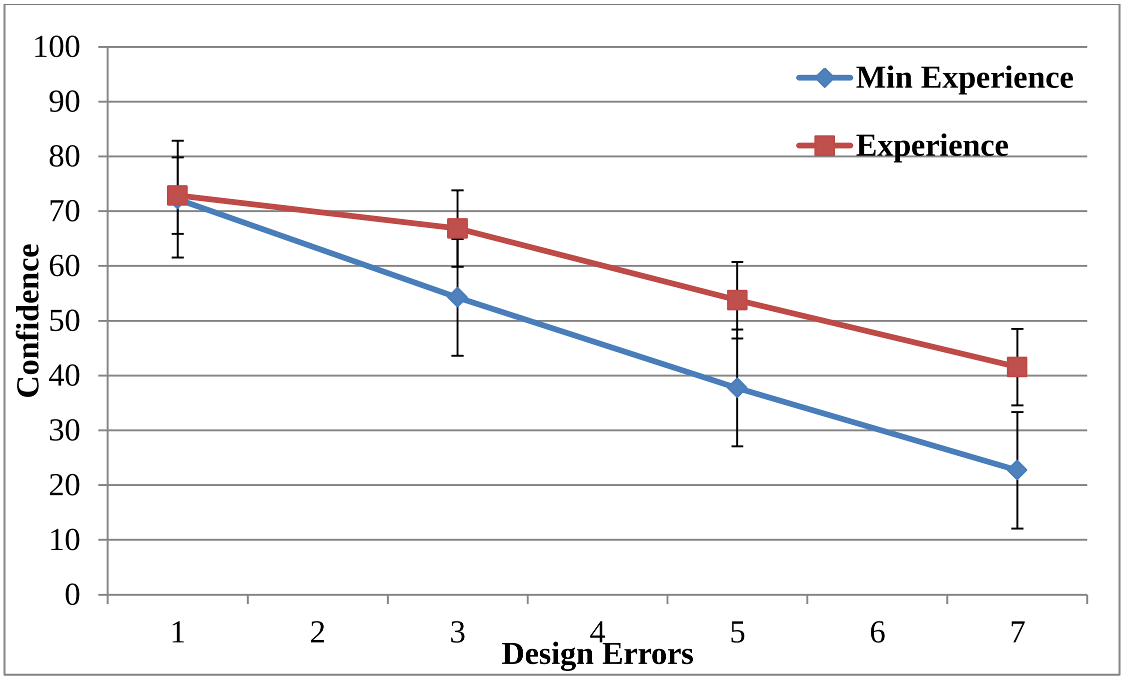


Figure 16: Experience (Domain Specialists) vs. Minimum Experience (Domain Generalists & Non-Domain Generalists)

Table 17: Confidence level for Experience vs. Minimum Experience (Combined - with and without controls)

# of Design Errors	Domain Specialists	Domain Generalists & Non-Domain Generalists	P-Value
1	72.86	72.2	0.9111
3	66.84	54.25	0.0362
5	53.76	37.74	0.0078
7	41.54	22.7	0.0019

4.3.3.2 Experience (Domain Specialists) vs. Minimum Experience (Domain Generalists & Non-Domain Generalists) – With controls

Figure 17 indicates the downward trend in the confidence level for the generalists and specialists as the design error increases. The confidence level for both the populations indicated is a combination of design errors with controls. For a single design error (Table 18), there is insufficient evidence to conclude there is a difference in the mean confidence among the generalists and the specialists (0.6006), indicating that the confidence level is the same. During assessment of the system performance for a single design error with control, generalists and specialists are considered comparable. When the system is assessed for a single design error with control, the experience of the individual does not matter.

For three design errors (Table 18), at a 5% level of significance there is insufficient evidence to conclude there is a difference in the mean confidence among the generalists and the specialists (0.1328), indicating that the confidence levels are comparable. During assessment of the system performance for three design errors, generalists and specialists are considered comparable. When the system is assessed for three design errors with control, the experience of the individual does not matter.

For five design errors and seven design errors, at a 5% level of significance there is sufficient evidence to conclude there is a difference in the mean confidence among the generalists and specialists (p=0.0191 & 0.0019 respectively), indicating that the confidence level is not the same. During assessment of the system performance for five design errors with control, care should be taken during decision making between an individual with experience and an individual with minimum experience.

Results suggest that when considering multiple design errors with controls, care should be taken during decision making for multiple failure modes as the experience of the individual does matter. *Specialists (maximum experience) were more optimistic about the performance of the system beyond five design errors. The difference was larger beyond five design errors since the non-domain generalists had an inclination point at three design errors.*

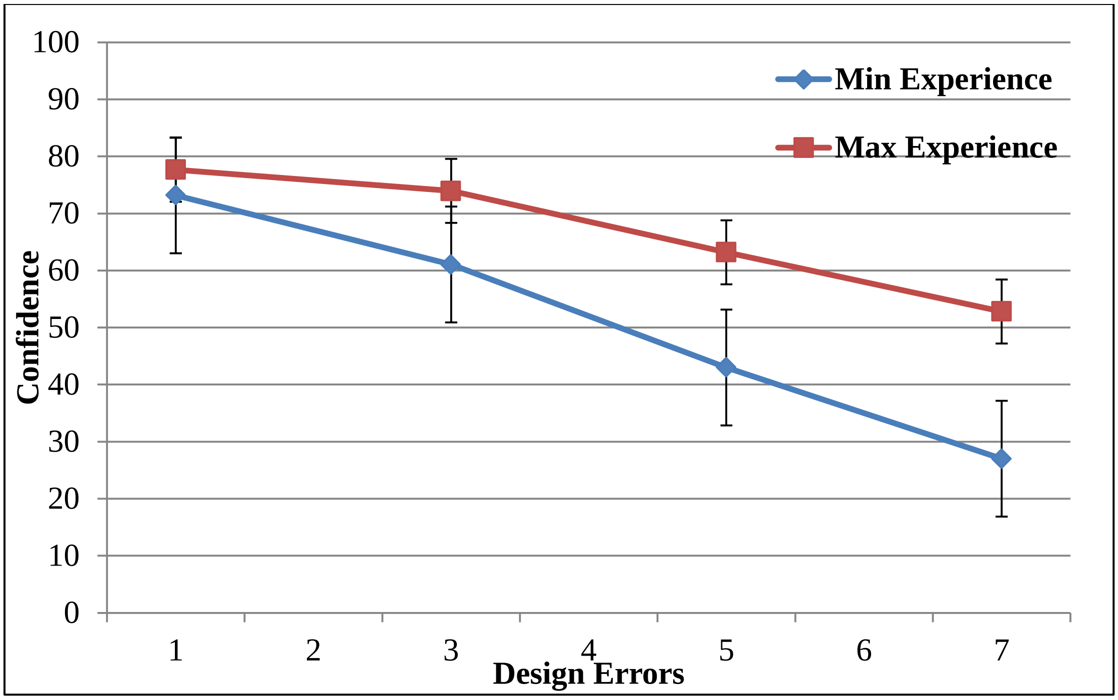


Figure 17: Experience (Domain Specialists) vs. Minimum Experience (Domain Generalists & Non-domain Generalists) - With Controls

Table 18: Confidence level of Domain Specialists vs. Domain Generalists & Non-Domain Generalists

# of Design Errors	Domain Specialists	Domain Generalists & Non-Domain Generalists	P-Value
1	77.66	73.18	0.6006
3	73.94	61.05	0.1328
5	63.19	43	0.0191
7	52.83	27	0.0028

4.3.3.3 Experience (Domain Specialists) vs. Minimum Experience (Domain Generalists & Non-Domain Generalists) – Without controls

Figure 18 indicates the downward trend in the confidence level for the generalists and specialists as the design error increases. The confidence level for both the populations indicated is a combination of design errors with controls. For a single design error (Table 19), there is insufficient evidence to conclude there is a difference in the mean confidence among the generalists and the specialists (0.7072), indicating that the confidence level is the same. For multiple design errors (Figure 18), at a 5% level of significance there is insufficient evidence to conclude there is a difference in the mean confidence among the generalists and the specialists (Table 19), indicating that the confidence levels are comparable.

Results suggest that when considering multiple design errors without controls, the estimation between the specialists and generalists are considered comparable. Specialists (maximum experience) were more optimistic about the performance of the system when controls are introduced. However, when controls are not introduced, the estimation between the two populations is comparable. *The difference between the two populations is the understanding of the use of controls. With the engineering experience the specialists could understand better, the use of controls and how this can influence the system performance. For a positive outcome, the decision of the generalists can be trusted, and for a negative outcome, the decision of the specialists can be trusted.*

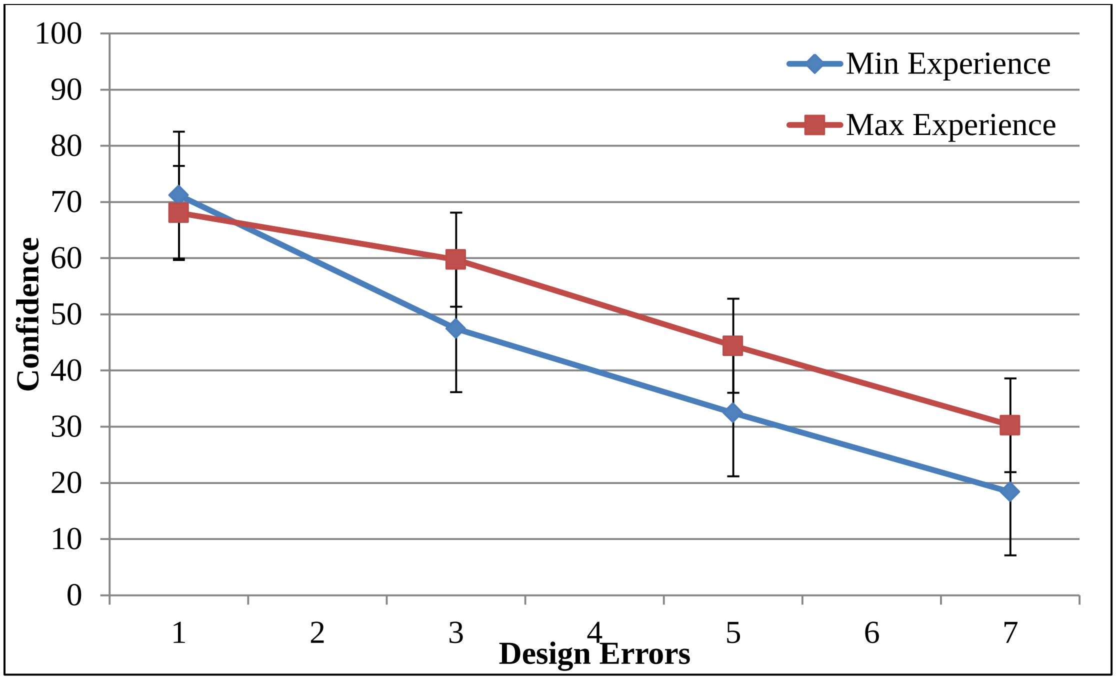


Figure 18: Experience (Domain Specialists) vs. Minimum Experience (Domain Generalists & Non-Domain Generalists) – Without controls

Table 19: Confidence level between Experiences (Domain Specialists) vs. Minimum Experience (Domain Generalists & Non-Domain Generalists) - Without controls

Design Errors	Domain Specialists	Domain Generalists & Non-Domain Generalists	P-Value
1	68.07	71.21	0.7072
3	59.74	47.44	0.1422
5	44.4	32.46	0.1544
7	30.26	18.4	0.1571

4.3.4 Engineering (Domain Specialists & Domain Generalists) vs. Non-Engineering (Non-Domain Generalists)

A comparison is made between the domain specialists and domain generalists against the non-domain generalists, to determine if the engineering experience of the individual makes a difference while estimating the system performance. Figure 19 indicates the downward trend in the confidence level for both the populations in estimating the system performance as the design error increases. The confidence level for both the populations indicated is a combination of design errors with and without controls.

4.3.4.1 Engineering (Domain Specialists & Domain Generalists) vs. Non-Engineering (Non-Domain Generalists) – Combined (with and without controls)

For a single design error and multiple design errors (Figure 19), there is insufficient evidence to conclude there is a difference in the mean confidence among the populations (Table 20), indicating that the confidence level is the same. The results indicate the comparable confidence level for multiple design errors when the domain specialists are combined with the domain generalists and compared against the non-domain generalists. The engineering knowledge of the domain specialists gained over a period of 3-4 years differentiates them from the domain generalists and non-domain generalists.

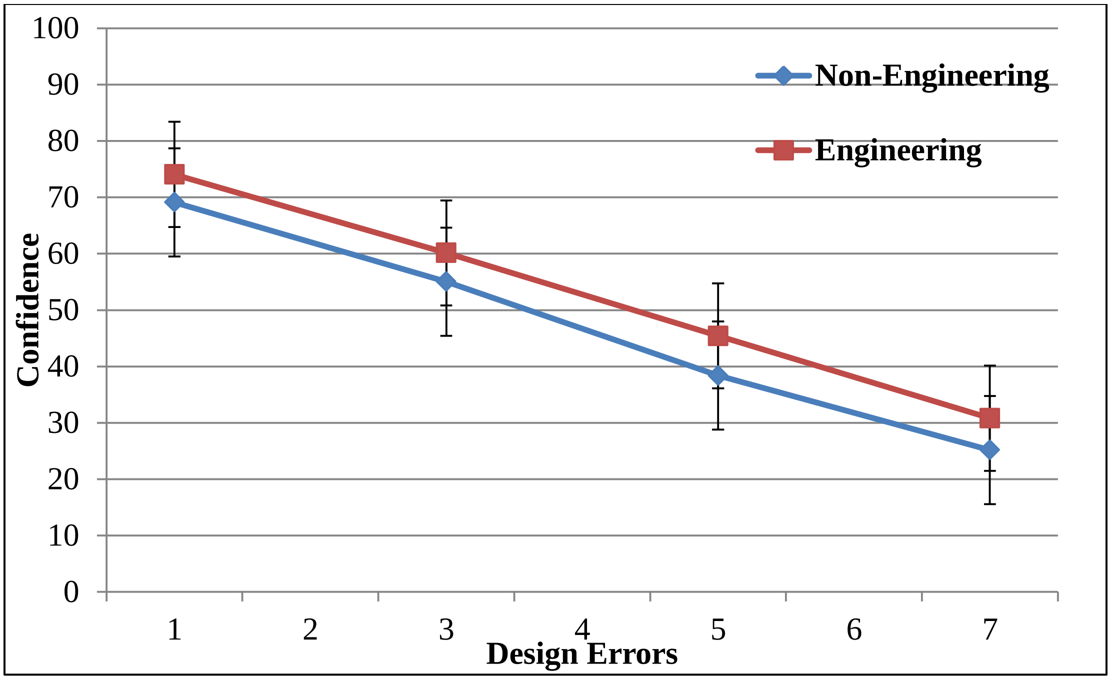


Figure 19: Engineering (Domain Generalists & Domain Specialists) vs. Non-Engineering (Non-Domain Generals)

Table 20: Confidence levels for Engineering (Domain Generals & Domain Specialists) vs. Non-Engineering (Non-Domain Generals)

Design Errors	Domain Specialists & Domain Generalists	Non-Domain Generalists	P-Value
1	74.06	69.13	0.3033
3	60.15	55.04	0.2870
5	45.43	38.38	0.1407
7	30.83	25.19	0.2358

4.3.4.2 Engineering (Domain Specialists & Domain Generalists) vs. Non-Engineering (Non-Domain Generalists) – With controls

Figure 20 indicates the downward trend in the confidence level for both the populations in estimating the system performance as the design error increases. The confidence level for both the populations indicated is for design errors with controls. For a single design error and multiple design errors (Figure 20), there is insufficient evidence to conclude there is a difference in the mean confidence among the populations (Table 21), indicating that the confidence level is the same.

The results indicate the comparable confidence level for multiple design errors when the domain specialists are combined with the domain generals and compared against the non-domain generalists. *The confidence level at three design errors is equivalent due to the inclination point at three design errors for the non-domain generalists. The learning of engineering knowledge of the domain specialists over a period of 3-4 years enables them to better understand the use of controls.*

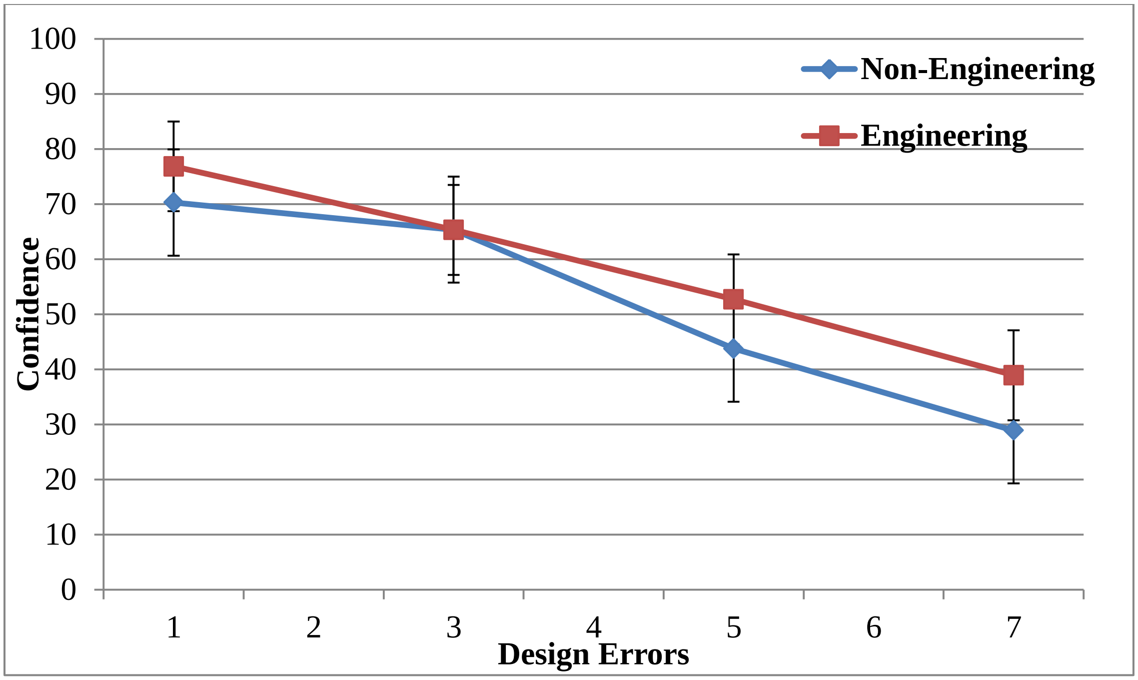


Figure 20: Engineering vs. Non-Engineering – With Controls

Table 21: Confidence level for Engineering vs. Non-Engineering – With Controls

Design Errors	Domain Specialists & Domain Generalists	Non-Domain Generalists	P-Value
1	76.86	70.3	0.3392
3	65.34	65.37	0.9957
5	52.71	43.78	0.1937
7	38.94	28.93	0.1455

4.3.4.3 Engineering (Domain Specialists & Domain Generalists) vs. Non-Engineering (Non-Domain Generalists) – Without controls

Figure 21 indicates the downward trend in the confidence level for both the populations in estimating the system performance as the design error increases. The confidence level for both the populations indicated is for design errors with controls. For a single design error and multiple design errors (Figure 21), there is insufficient evidence to conclude there is a difference in the mean confidence among the populations (Table 22), indicating that the confidence level is the same.

The results indicate the comparable confidence level for multiple design errors when the domain specialists are combined with the domain generalists and compared against the non-domain generalists. *The engineering knowledge of the domain specialists over a period of 3-4 years enables them to better understand the implications of the failure mode.*

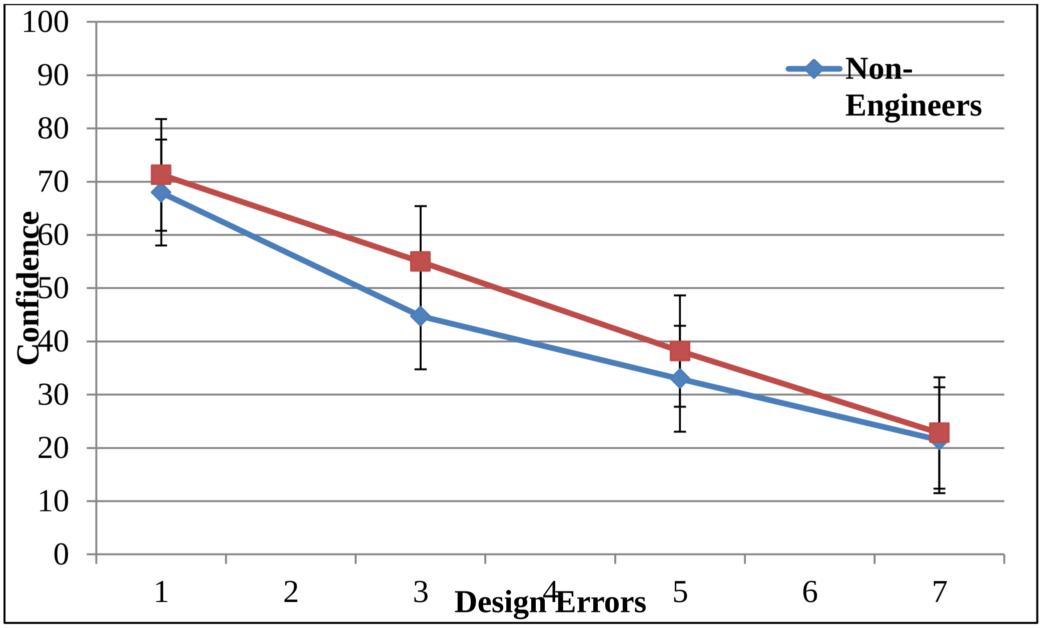


Figure 21: Engineers vs. Non-Domain Engineers - Without controls

Table 22: Confidence level for Engineers vs. Non-Domain Generals

Design Errors	Domain Specialists & Domain Generalists	Non-Domain Generalists	P-Value
1	71.27	67.96	0.6210
3	54.95	44.71	0.1260
5	38.18	32.97	0.4360
7	22.81	21.45	0.8390

4.4 Major Results and Takeaways

Table 23 is a summary of the critical results obtained from the analysis conducted in this research. The table indicates the p-value, and takeaways of key comparisons between the domain populations.

DS: Domain Specialists; DG: Domain Generalists; NDG: Non-Domain Generalists

Table 23: Summary of Analysis and Results

Design Errors	Comparison	P-Value	Takeaway
Results on Domain and Design Error Interaction			
1	DS vs. DG vs. NDG	0.3456	It does not matter who the decision maker is since the assessment of the system performance for all domain populations are comparable.
3	DS vs. DG vs.	0.0793	When the number of design errors is three or

	NDG		more, care should be taken during decision making since the difference in the assessment of the system performance among the domain populations is evident.
5	DS vs. DG vs. NDG	0.0198	As the number of failure modes considered at a given time increases, the difference in assessment is larger.
7	DS vs. DG vs. NDG	0.0002	The trend in confidence level suggests that the domain specialists are more optimistic in predicting the performance of the system.
Percent Decrease in Confidence			
Rate of decrease in assessment for DG & NGD		0.0002	The plummet in the confidence level of the non-domain generalists and domain generalists indicate they are pessimistic about the performance of the system while considering multiple failure modes at the same time.
Rate of decrease in assessment for DS		0.0301	The rate of decrease in confidence is the least for the domain specialists and they remain optimistic about the system performance.
Comparison between the three domain populations for rate of			If the specialist gives a negative estimation of the system success, they can be trusted. In other

decrease in confidence level			words, the domain generalists can be trusted when they give a positive estimation of the system success and cannot be trusted when they give a negative response.
Effect of Controls on all Domain populations			
1	Combination of domain populations	0.3348	Providing control for a single error does not significantly improve the confidence level in predicting the system performance.
3, 5, 7	Combination of domain populations	0.0049	Providing controls for multiple errors increased the confidence level significantly in predicting the system performance.
1	DG	0.6794	It does not matter if control is being used for a single design error since the confidence level in assessing the performance of the system is comparable.
3, 5, 7	DG	0.0089	For multiple design errors, the introduction of controls increased the confidence level for the domain generalists in estimating the system performance. However, the use of controls did not affect the sense of pessimism in estimating the system success. The reason for this is the lack of engineering expertise in understanding

			the implications of using controls.
1	NDG	0.7558	It does not matter if control is being used for a single design error since the confidence level in assessing the performance of the system is comparable.
3	NDG	0.0066	Point of transition was achieved, beyond three design errors, the confidence level decreased drastically.
5 & 7	NDG	0.3201	The trend in the confidence level for both, domain generalists and non-domain generalists can be considered comparable since the domain generalists have only 5-6 months of engineering knowledge and the non-domain generalists have almost no engineering knowledge over their course of study.
1 & 3	DS	0.3916	The knowledge or domain awareness of the specialists enables them to make an optimistic estimation of the system performance for multiple errors
5 & 7	DS	0.0448	The shape of the trends suggests that the domain generalists and specialists have near linear relationships with the number of increasing

			design errors
Pairwise Comparisons			
DG vs. NDG		There is no significant difference between the two domain populations when design errors are considered with controls, without controls, or combined. This result is validated by the fact that the domain generalists have an engineering experience of less than six months. Hence the domain generalists and the non-domain generalists can be considered to be pessimists in predicting the system success. Their decision can be trusted if they provide a positive outcome on the performance of the system	
DG vs. DS		The specialists are optimistic about the performance of the system, hence, their decision for multiple failure modes can be trusted if the decision has a negative outcome. The domain specialists are pessimistic about the system performance and their decision can be trusted if their decision has a positive outcome. There is a possibility that beyond seven design errors, the confidence level in estimation of system performance could have a point of inclination, beyond which the confidence level decreases drastically	
DS vs. DG & NDG		The decision made by an individual with maximum	

	experience on the system performance for multiple design errors can be trusted if the estimation has a negative outcome. In case of a positive outcome, the domain generalists and non-domain generalists can be trusted. Specialists (maximum experience) were more optimistic about the performance of the system beyond five design errors. The difference was larger beyond five design errors since the non-domain generalists had an inclination point at three design errors. The difference between the two populations is the understanding of the use of controls. With the engineering experience the specialists could understand better, the use of controls and how this can influence the system performance.
DS & DG vs. NDG	The confidence level at three design errors is equivalent due to the inclination point at three design errors for the non-domain generalists. The learning of engineering knowledge of the domain specialists over a period of 3-4 years enables them to better understand the use of controls and the failure modes

CHAPTER FIVE: CONCLUSIONS

Design reviews are typically used for: 1) Identifying errors, 2) Assessing the impact of the errors, and 3) Suggesting solutions for the errors. The study in this research focuses on understanding the second issue as it relates to the number of errors considered, existence of controls, and the level of domain familiarity of the assessor. The research presents a study on how estimations are made in design reviews between domain generalists, domain specialists, and non-domain generalists. The findings in the research helps in evaluating the importance of methods used in engineering design and how they affect quality estimations of different types of designers on system performance. Non domain generalists, domain generalists, and domain specialists are provided a set of design errors and asked to estimate the likelihood that the system would successfully achieve the objectives.

5.1 Estimating System Performance

During FMEA, investigation is conducted by considering single errors at a time, results in this research suggest that when considering one design error at a time, it does not matter if the decision maker is a domain generalist or a domain specialist, they can be considered equivalent. However, while considering multiple design errors at a given time, there is a significant difference in estimating the performance. The mean confidence in estimating the system performance for the domain specialists is higher than domain and non-domain generalists for multiple errors. The difference is larger as the number of errors increases. The rate of decrease in confidence is faster for domain generalists and the rate of decrease is least for domain specialists. Considering the experience and knowledge in the domain, suggests that the domain specialists are more optimistic of the system performance based on their deeper understanding of the functionality of the system. The shape of the trends suggests that the domain generalists and specialists have near linear relationships with the number of increasing design errors. The non-domain generalists (psychology class students) do not show such a linear trend. The Non-domain generalists achieved a point of transition beyond three domain errors, beyond which the confidence level decreased drastically. There is a possibility that beyond seven design errors, there is a possibility that the specialists' estimation could have a point of inclination, beyond which the confidence level decreases drastically. The trend in the confidence level for both, domain generalists and non-domain generalists can be considered comparable since the domain generalists have only 5-6 months of engineering knowledge and the non-domain generalists have almost no engineering knowledge over their course of study.

5.2 Use of Controls

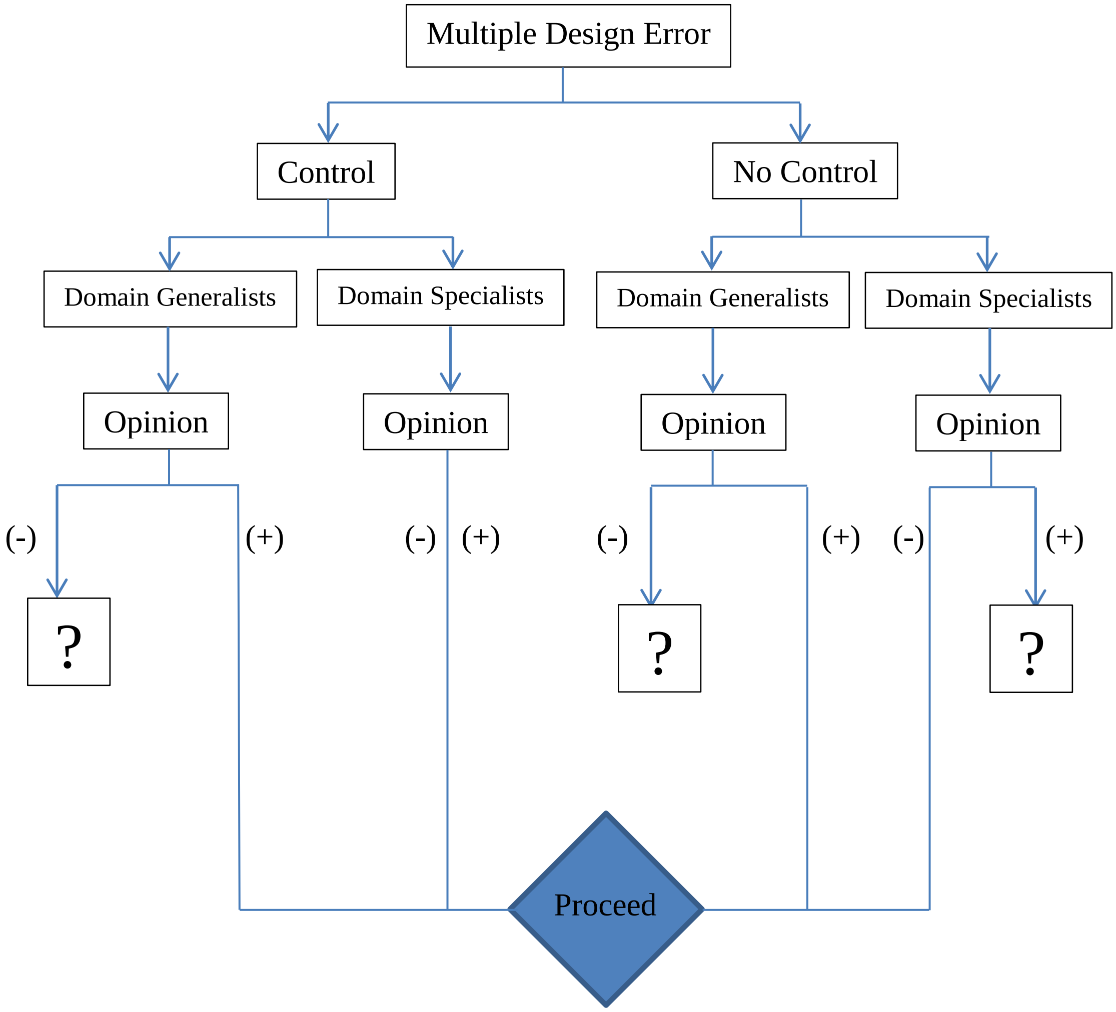
The increase in confidence level due to the use of control was evident while considering multiple design errors at the same time. Providing controls to multiple failure modes increased the confidence level significantly. However, the confidence level with the use of control had a downward trend. The use of controls proved to be more significant for the domain specialists as they understand better the implications of the use of controls to reduce the risk or prevent the errors. The domain specialists should be able to conceptualize how the specific offered controls could be implemented and executed. The other generalist populations, however, do not have this contextual background on which to base these estimations. Hence the use of controls did not affect the pessimistic estimation of the generalists. However, a positive decision about success by the generalists can be encouraged.

5.3 Decision Making

Domain specialists showed a more optimistic point of view of system success (higher confidence) in making the decision with design errors. This suggests that, if the specialist gives a negative estimation of the system success, they can be trusted. The plummet in the confidence level of the non-domain generalists and domain generalists indicate they are pessimistic about the performance of the system while considering multiple failure modes at the same time. The domain generalists can be trusted when they give a positive estimation of the system success and cannot be trusted when they give a negative response.

It is necessary to allow engineers with the appropriate expertise to make judgments during the design review or product development stage. Since the opinions differ, design review teams are facing challenges to make the right decision with confidence. Decisions related to the failure of a system should be given most importance when made by a domain specialist. In other words, the decisions related to the success of a system can be trusted when made by a domain generalist. Decisions made on a system having failure modes with associated controls should be given most importance when made by a domain specialist.

Figure 22: Decision Flow Diagram for Multiple Design Errors



5.4 Further Investigation

The study tested the hypothesis that the design errors impact decision making in design reviews. This hypothesis is proven and other findings were also made. Additionally, the confidence level for all populations decreased significantly as the number of design errors increased. Further, the confidence decreases less when solutions (controls) are provided to mitigate the errors. The domain specialists are optimistic in predicting the performance of the system than domain generalists and non-domain generalists as the design errors increase.

5.4.1 Transition point

It is not clear whether there is a plateau effect at higher numbers of design errors considered simultaneously, such as at 25 errors. Perhaps there is an asymptotic level at which adding new design errors does not introduce any new perceived degradation of the design performance. Non-domain generalists achieved a transition point beyond three domain errors. The transition point for the domain specialists could be achieved for higher number of design errors. Additional trends could be planned to explore the assessment trends for the domain specialists.

5.4.2 SAS Code

The study in this research compares a dependent variable (confidence level) with three independent variables (design errors, controls and domain population) and analysis is conducted to derive significant statistics between the variables. The SAS code developed for this research can be used to:

1. Derive significant statistics to differentiate trends between curves. In this research the difference in the weighted average of the domain populations is indicated for varying number of design errors
2. Derive significant statistics when more than two independent variables are involved. The interaction between two independent variables can be studied by blocking the effect of the third independent variable. Example: The study in this research involves three independent variables: Controls, design errors, and domain population. The interaction between Controls and design errors can be studied by blocking the influence of the domain population.
3. Sort data according to the variable of interest. In this research, the data is sorted according to domain populations, controls and number of design errors.
4. Compare one independent variable with the average of the other two independent variables. Example: In this research a comparison is made between the domain specialists vs. domain generalists & non-domain generalists. The confidence level of the domain specialist is compared with the average of the confidence levels of the domain generalists and non-domain generalists.

5.4.3 Industry Level Experiment

The experiment layout in this research involved individual estimation of the performance prediction. It would be interesting to see how the assessment of the system performance would differ when individuals are grouped within the same domain or between domains. The diagonal elements in Table 24 is hypothesized to behave like individual assessors, the off diagonal elements can be predicted by statistical estimates. The results from this can indicate the role played by group dynamics in making decisions. Proposed experiment layout for grouped assessment:

DG: Domain Generalist; NDG: Non-Domain Generalist; DS: Domain Specialist

Table 24: Group Assessment

	DG	NDG	DS
DG	DG+DG	DG+NDG	DG+DS
NDG	-	NDG+NDG	NDG+DS
DS	-	-	DS+DS

The participants in this study involved students having the most engineering experience (graduate students), least engineering experience (general engineers), and non-engineering (no engineering experience). Replication of this study can be carried out in an industry to differentiate between entry level engineers, interns, managers and engineers from various domains having different levels of engineering experience. It would be interesting to see the difference in the estimation of entry level engineers, who have graduated with a bachelors' degree against graduate level engineering students. Theoretically, the assessment should be comparable since they have equal engineering experience. Hence the difference would be industry level experience and graduate level experience.

5.4.4 Differential Factor

The study indicated a difference in estimation of system success for individuals at varying level of expertise. By comparing the estimates, a factor can be identified which can be used to determine the assessments made by individuals at varying level of expertise. Example: The confidence level in predicting the system performance by a domain generalist is 55%. If a domain specialist would give his assessment, it can be calculated by 55%*differential factor. By achieving this, the trust imparted by a domain specialist can be found out by using an assessment made by a domain generalist.

5.4.5 Assessment of performance prediction by Junior Engineers

The participants in this research involved Graduate Mechanical Engineering students (5 years of engineering experience) and General Engineering students (6 months of engineering experience). An experiment on Junior Mechanical Engineering students (3 years of engineering experience) can be conducted to confirm if the assessment is the average of the generalists and specialists.

5.4.6 Further questions that can be posed based on this research

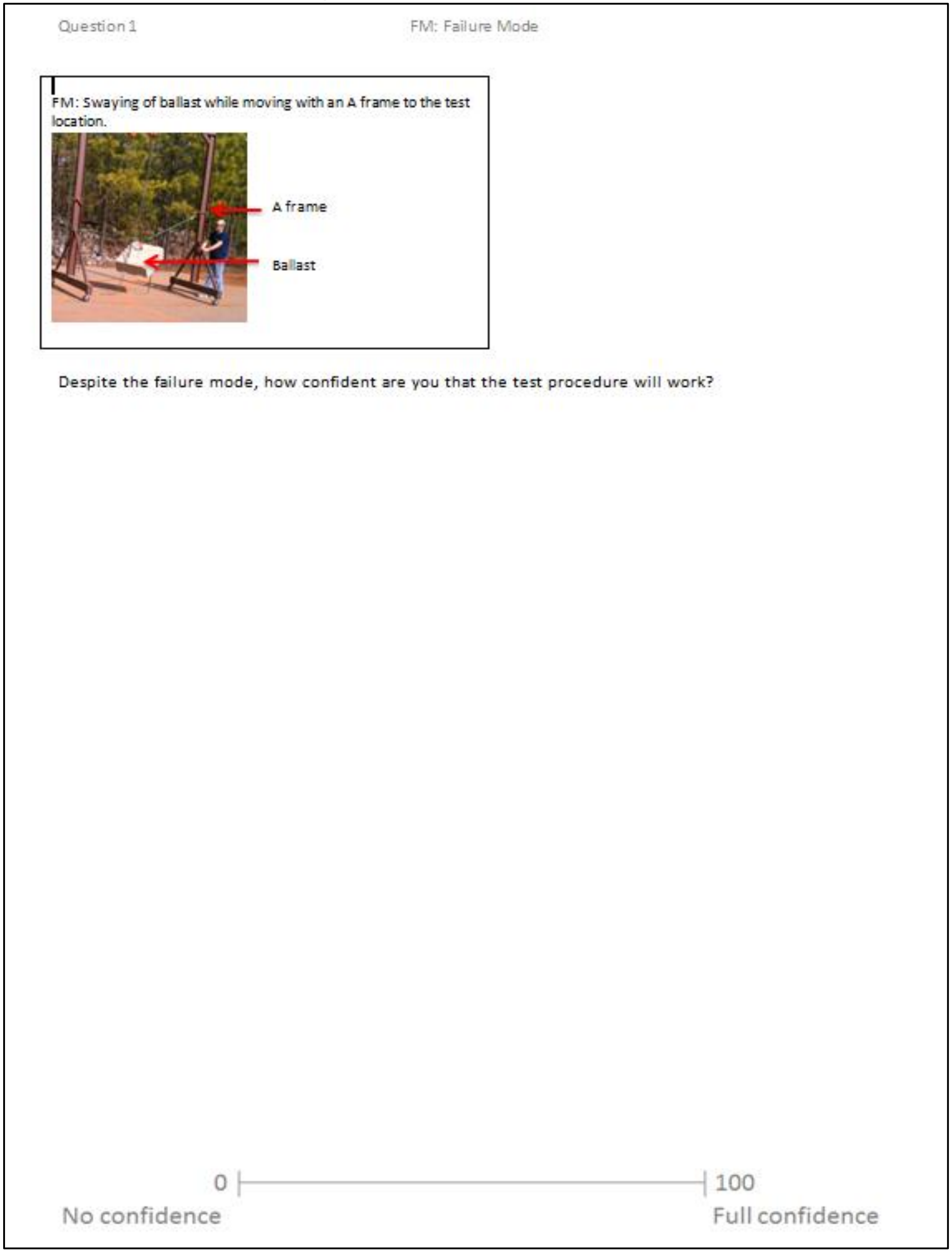
1. What is the maximum number of errors that can be considered at a given time before the confidence level in predicting system performance for the domain specialists reaches saturation?
2. Is the estimation of system success by Junior Engineering students, an average of the estimation for Graduate Engineering students and General Engineering Students?
3. What is the impact of group assessments in decision making?

APPENDICES

Appendix A: User Study Packets

User study packets are used to collect data from the participants as explained in section 3.3.5. The participants were divided into two groups without regard to gender, race, or personality. The first group was provided design errors without controls and the second group was provided design errors with controls. The failure mode is represented at the top of the page and based on the number of failure modes presented, the participants answer the question *"Despite the failure mode, how confident are you that the procedure will work?"* by marking a slash line on the linear scale. The failure mode is represented using a picture as well as in a textural description. The confidence level is measured using a continuous scale from the point 0 until the point of intersection of the slash mark on the linear scale. The data is collected all analyzed using SAS (Statistical Analysis System).

Experiment group 1: Failure modes without control



Question 1

FM: Failure Mode

FM: Swaying of ballast while moving with an A frame to the test location.

Despite the failure mode, how confident are you that the test procedure will work?

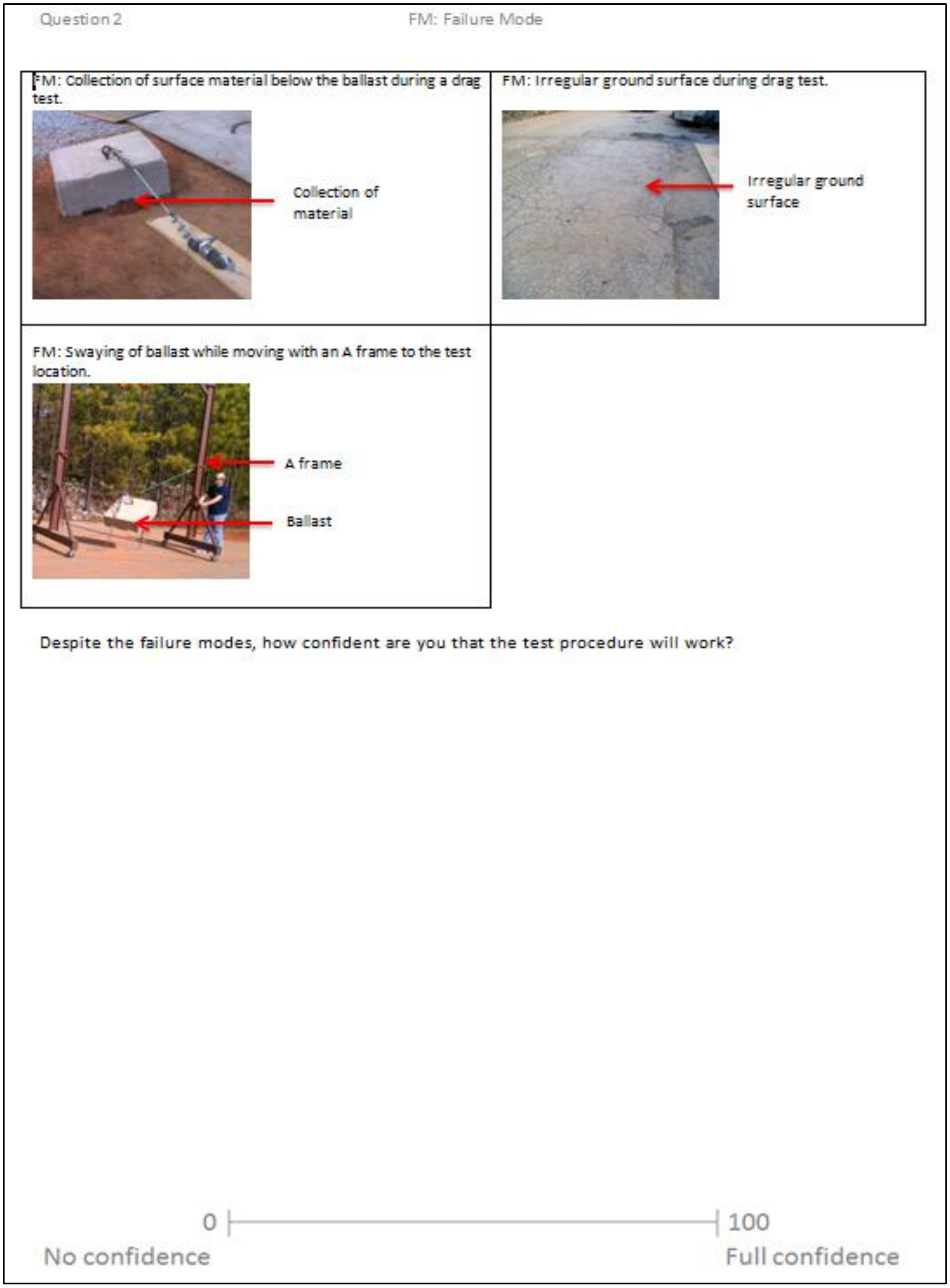
0

100

No confidence

Full confidence

Question 1



Question 2

FM: Failure Mode

FM: Collection of surface material below the ballast during a drag test.

Collection of material

FM: Irregular ground surface during drag test.

Irregular ground surface

FM: Swaying of ballast while moving with an A frame to the test location.

A frame

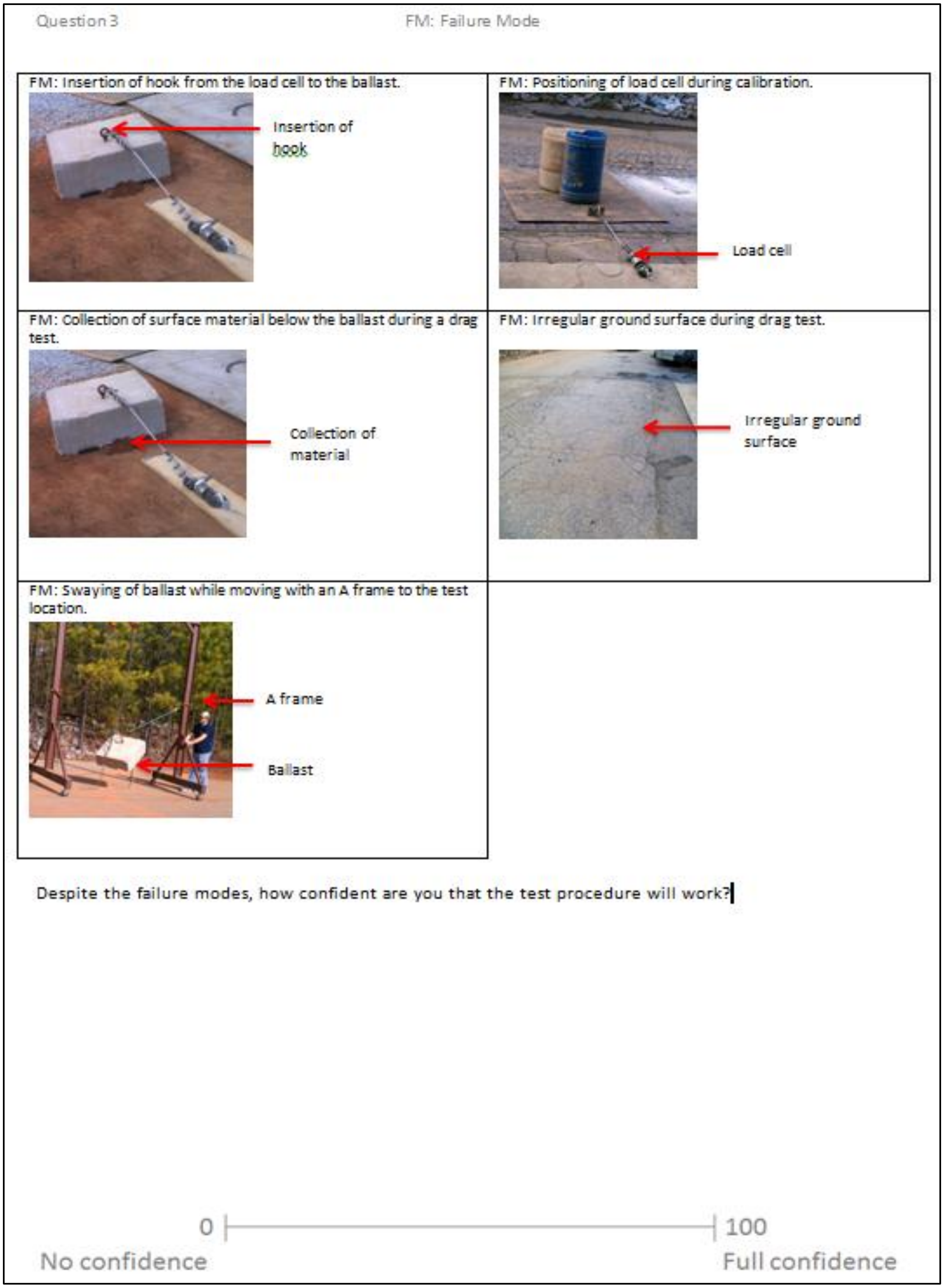
Ballast

Despite the failure modes, how confident are you that the test procedure will work?

0 No confidence

100 Full confidence

Question 2



Question 3

FM: Failure Mode

FM: Insertion of hook from the load cell to the ballast.

FM: Positioning of load cell during calibration.

FM: Collection of surface material below the ballast during a drag test.

FM: Irregular ground surface during drag test.

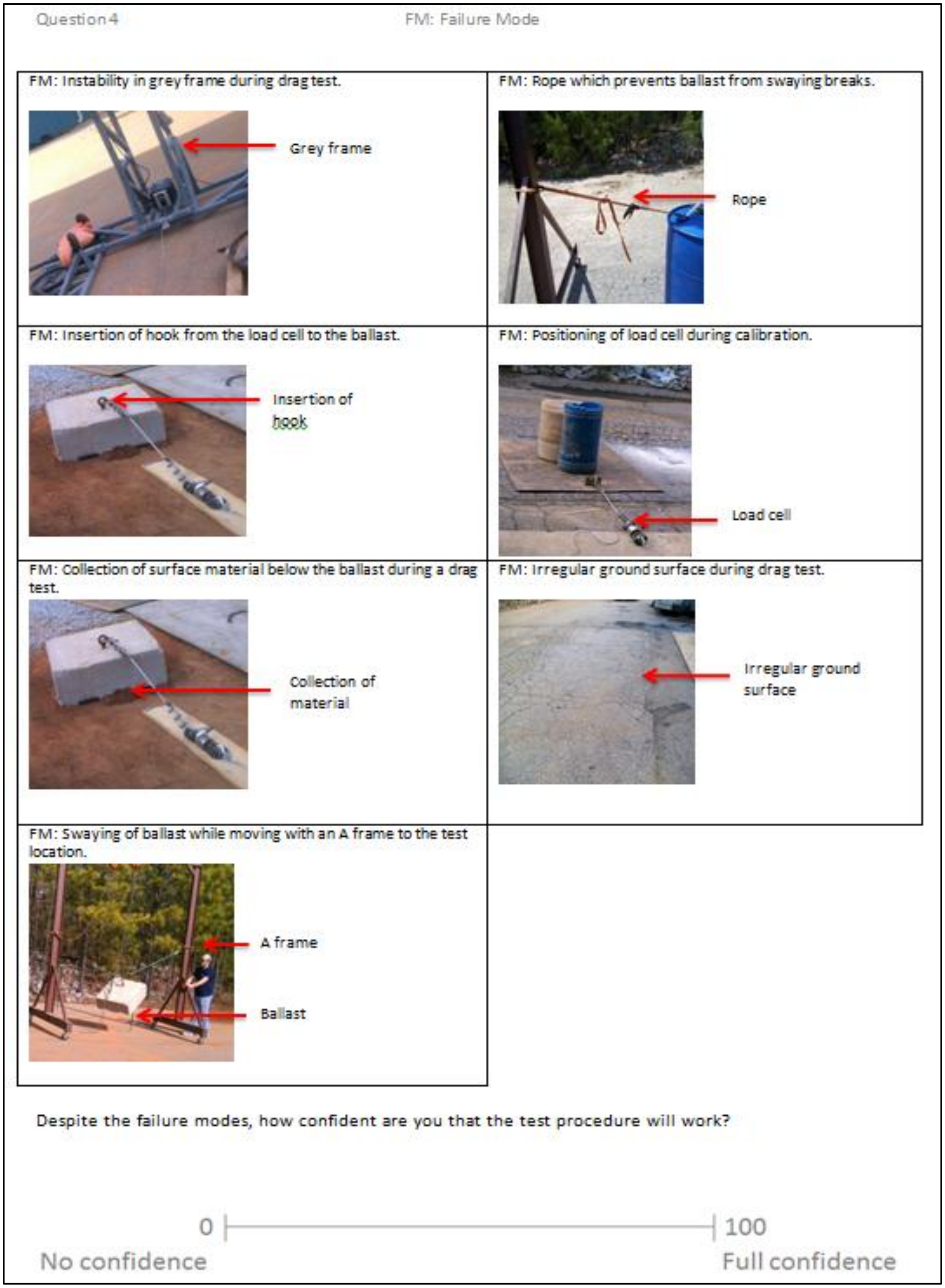
FM: Swaying of ballast while moving with an A frame to the test location.

Despite the failure modes, how confident are you that the test procedure will work?

0 No confidence

100 Full confidence

Question 3



Question 4

FM: Failure Mode

FM: Instability in grey frame during drag test.

Grey frame

FM: Rope which prevents ballast from swaying breaks.

Rope

FM: Insertion of hook from the load cell to the ballast.

Insertion of hook

FM: Positioning of load cell during calibration.

Load cell

FM: Collection of surface material below the ballast during a drag test.

Collection of material

FM: Irregular ground surface during drag test.

Irregular ground surface

FM: Swaying of ballast while moving with an A frame to the test location.

A frame

Ballast

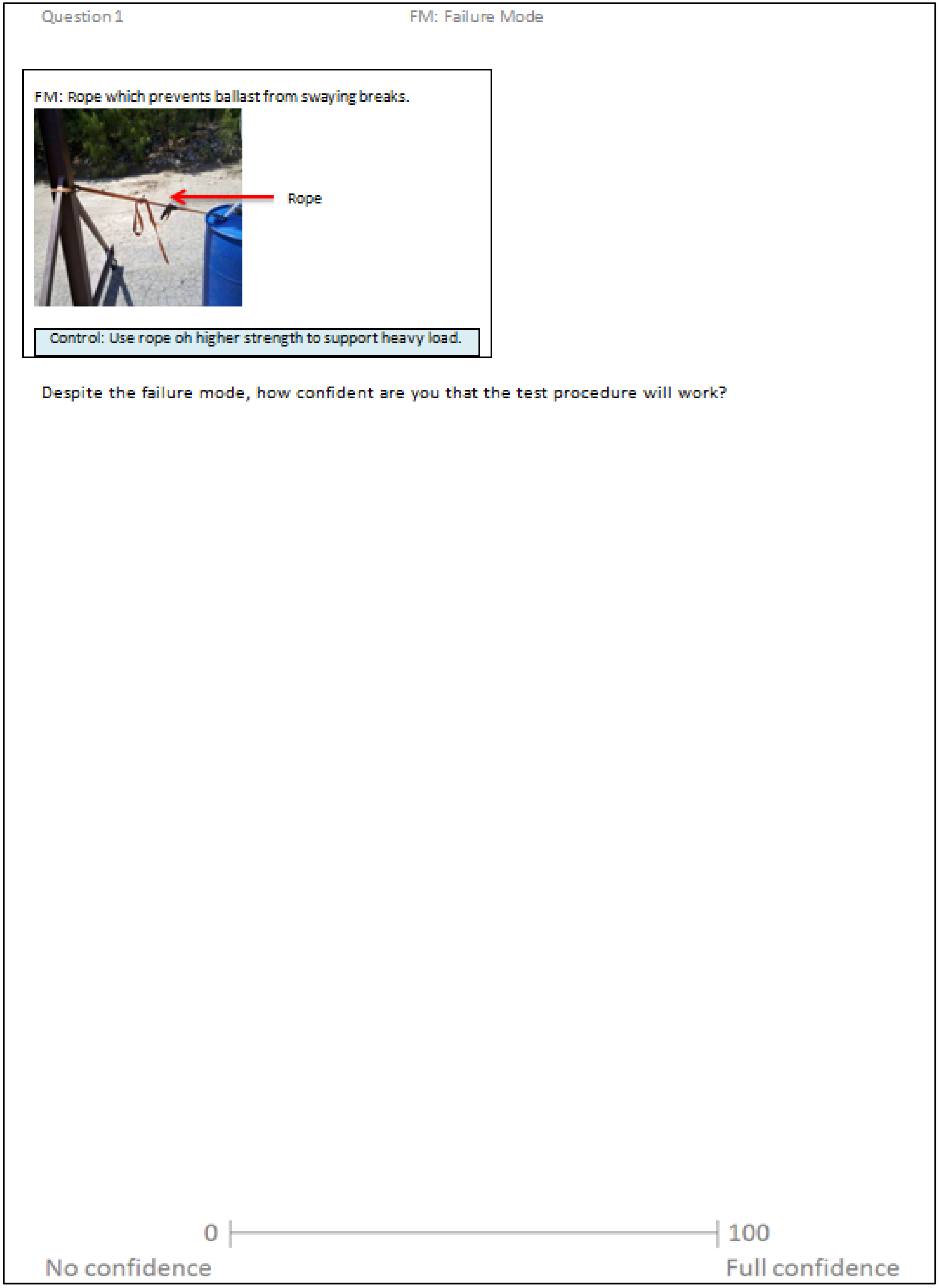
Despite the failure modes, how confident are you that the test procedure will work?

0 No confidence

100 Full confidence

Question 4

Experiment group 2: Failure modes with control



Question 1

FM: Failure Mode

FM: Rope which prevents ballast from swaying breaks.

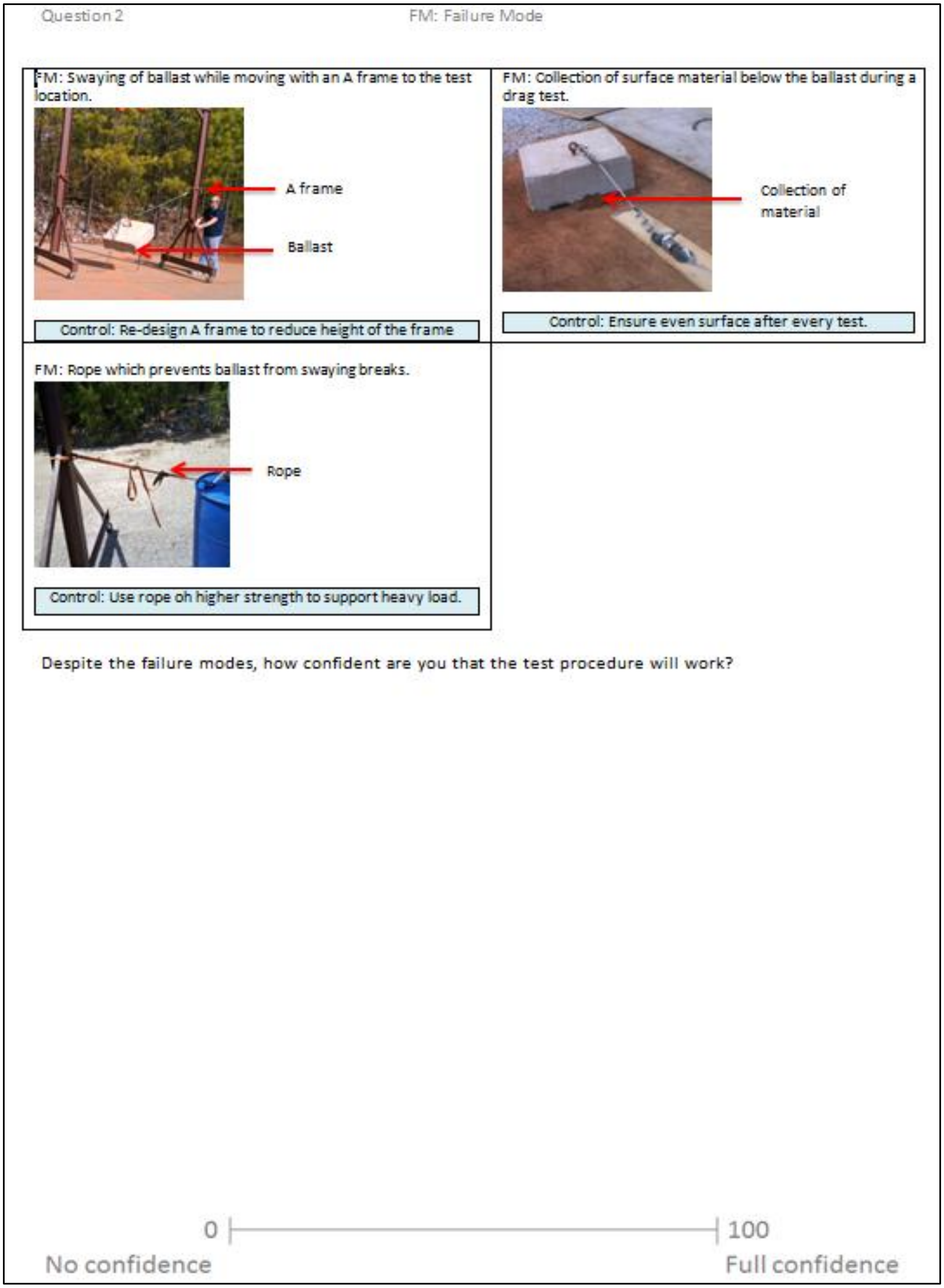
Control: Use rope oh higher strength to support heavy load.

Despite the failure mode, how confident are you that the test procedure will work?

0 No confidence

100 Full confidence

Question 1



FM: Swaying of ballast while moving with an A frame to the test location.

Control: Re-design A frame to reduce height of the frame

FM: Collection of surface material below the ballast during a drag test.

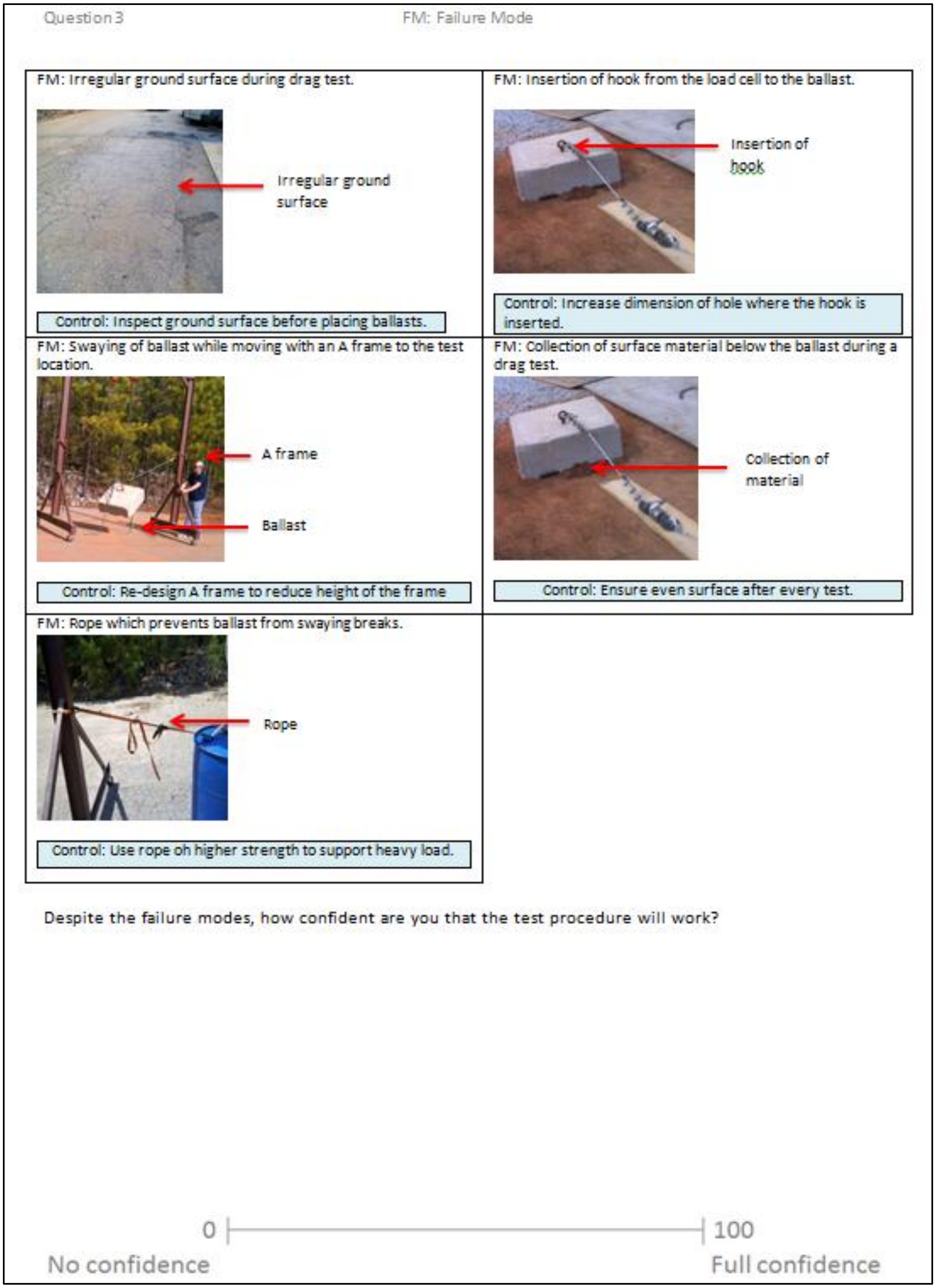
Control: Ensure even surface after every test.

FM: Rope which prevents ballast from swaying breaks.

Control: Use rope oh higher strength to support heavy load.

Despite the failure modes, how confident are you that the test procedure will work?

Question 2



Question 3

FM: Failure Mode

FM: Irregular ground surface during drag test.

Control: Inspect ground surface before placing ballasts.

FM: Insertion of hook from the load cell to the ballast.

Control: Increase dimension of hole where the hook is inserted.

FM: Swaying of ballast while moving with an A frame to the test location.

Control: Re-design A frame to reduce height of the frame

FM: Collection of surface material below the ballast during a drag test.

Control: Ensure even surface after every test.

FM: Rope which prevents ballast from swaying breaks.

Control: Use rope oh higher strength to support heavy load.

Despite the failure modes, how confident are you that the test procedure will work?

Question 3

Question 4

FM: Failure Mode

FM: Positioning of load cell during calibration.

Control: Check load value after every trial during calibration.

FM: Instability in grey frame during drag test.

Control: Place rubber mats below the frame to increase friction.

FM: Irregular ground surface during drag test.

Control: Inspect ground surface before placing ballasts.

FM: Insertion of hook from the load cell to the ballast.

Control: Increase dimension of hole where the hook is inserted.

FM: Swaying of ballast while moving with an A frame to the test location.

Control: Re-design A frame to reduce height of the frame

FM: Collection of surface material below the ballast during a drag test.

Control: Ensure even surface after every test.

FM: Rope which prevents ballast from swaying breaks.

Control: Use rope of higher strength to support heavy load.

Despite the failure modes, how confident are you that the test procedure will work?

Question 4

Appendix B: Data Collection

General Engineering Students (Domain Generalists)

Confidence Level for Modes with Control				
Student #	1	3	5	7
1	22.8	13	12	7.6
2	96.7	76.1	70.7	50
3	97.8	80	65.2	51.1
4	96.7	87	77.2	71.7
5	76.1	67.4	60	52.2
6	79.3	71.7	66.3	53.3
7	70.7	50.5	29	22.6
8	64.1	45.7	34.8	5.4
9	75	41.3	14.1	16.3
10	75	50	47.8	32.6
11	75	55.4	40.2	16.3
12	17.4	20.7	4.3	0
13	89.1	65.2	22.8	2.2
14	86	65.2	56.5	50
15	82.6	74	65.2	61
16	22.8	22.8	7.6	3.3
17	89.1	30.4	31.5	2.2
18	97.8	78.3	75	31.5
19	100	56.5	48.9	33.7
20	62	21.7	12	10.9
21	79.6	65.6	35.5	11.8
22	86	68.8	50.5	25.8
23	95.7	71	46.2	10.8
24	83.9	75.3	44.1	28
25	83.9	66.7	52.7	33.3
26	79.6	62.4	52.7	24.7
27	93.5	75.3	57	38.7
28	67.7	50.5	29	14
29	74	52	31	23
30	63.4	32.3	19.4	10.8
31	61.3	30.1	17.4	8.4

Confidence Level for Modes Without Control				
Student #	1	3	5	7
1	66.3	54.3	7.6	0
2	94.6	71.7	19.6	13
3	75	63.8	47.4	13.2
4	76.1	58.7	38	2.2
5	47.8	22.8	2.2	1.1
6	76.1	48.9	19.6	13
7	84.8	66.3	58.7	47.8
8	87	63	50	49
9	87	73.9	55.4	52.2
10	95.8	74	52.2	42.4
11	68.5	43.5	42.4	34.8
12	60	53.3	41.3	23.9
13	65.2	43.5	21.7	7.6
14	67.4	34.8	18.5	9.8
15	76.1	27.2	13	0
16	96.7	66.3	19.6	0
17	54.3	24	7.6	0
18	48.4	28	22.6	15.1
19	82.8	64.5	48.4	24.7
20	74.2	69.9	52.7	17.2
21	69.9	50.5	32.3	6.5
22	94.6	67.7	49.5	10.8
23	91.2	60.2	35.5	14
24	81.7	63.4	61.3	26.9
25	69.9	50.5	28	0
26	31.2	33.3	19.4	16.1
27	82.8	64.5	41.9	18.3
28	75.3	48.4	31.2	6.5
29	68.8	20.4	3.2	1.1
30	72	51.6	38.7	24.7
31	76.2	52.3	33.8	17.2

32	93.5	66.3	30.4	28
33	93.5	65.2	60.9	45.6
34	72	67	60	38
35	78.3	59.4	44.3	27.2
36	74.8	55.3	40.2	23.4
37	81.2	62.6	47.8	30.2
38	71.8	52.8	37.4	21.7
39	70	65.3	57.9	36
40	90	64.5	60	23.1
41	89.8	65.2	53.3	12
42	87	72.8	63	42.4
43	87	71.7	55.4	45.7
44	73.9	51.1	29.9	8.6
45	62.4	48.4	35.9	6.5
46	84.6	73.6	39.1	0
47	71.7	50	28.3	22.3
48	40.9	29	15.1	7.6
49	77.7	54.3	43.4	33
50	53.1	34	25	6.8
51	75.2	55.8	41.6	24.2
52	62.4	44.1	32.3	12.9

32	86.9	34.7	36.9	42.4
33	84.9	50.5	36.6	20.4
34	72.2	48.3	30.1	13.8
35	84.1	60.3	42.3	22.4
36	69.6	45.7	27.4	10.3
37	64.8	40.2	22.3	5.8
38	79.3	55.2	37.8	18.4
39	81.2	57.3	39.4	22.6
40	67.8	43.2	25.6	7.5
41	71.6	47.8	29.6	11.4
42	79.2	60.3	45.4	27.7
43	69.9	58.1	38.7	0
44	88.2	71	35.5	12.9
45	67.7	31.2	47.3	10.8
46	79.8	33	12.2	0
47	65.2	21.3	3.3	0
48	59.8	19.4	3.2	3.3
49	83	66.3	50.5	29
50	68	58.1	34.4	3.2

Psychology class students (Non- Domain Generalists)

Confidence level for Modes with Control				
Student #	1	3	5	7
1	87	79.6	77.4	76.3
2	52.7	40	20.4	9.6
3	74.2	65.6	23.7	5.4
4	75.3	38.7	10.8	5.4
5	66.7	50.5	30.1	13
6	87	86	73.11	43
7	82.8	64.5	25.8	14
8	88.2	77.4	66.7	43
9	40	42	24.7	13
10	72.5	67.5	45.9	31.2
11	68.2	63.2	41.6	26.7

Confidence level for modes without Control				
Student #	1	3	5	7
1	63.4	39.8	18.3	9.7
2	78.5	61.3	24.7	6.5
3	45.2	32.3	16.1	8.6
4	33.3	23.7	19.4	9.7
5	72	3.2	0	0
6	42	40.8	28	24.7
7	73.1	48.4	16.1	0
8	68.8	12.9	8.6	2.2
9	77.4	79	62.4	50.5
10	88.2	71	32.3	18.3
11	81.7	57	32.3	10.8

12	66.1	61.4	38.2	24.5
13	74.6	69.4	46.3	24.2
14	70.2	65.1	42.9	30
15	79.6	89.24	52.9	25.8
16	48.4	68.8	51.6	48.4
17	71	75.3	60.9	50
18	61.3	78.5	45.2	13
19	68.9	66.1	43.2	28.9
20	70	59.1	46.2	45.2

12	86	72	64.5	37.6
13	66.7	14	73.1	18.3
14	64.6	41.2	29.4	17.4
15	72.2	49.3	37.4	25.6
16	76.4	47.3	44	68.8
17	66.7	68	54.8	54

Graduate Mechanical Engineers (Domain Specialists)

Confidence level without Control				
Student #	1	3	5	7
1	92.4	89.1	73.9	56.5
2	43.5	32.6	61	25
3	67.4	62	64	62
4	20	37	21	16.3
5	52.2	28.3	19.6	9.8
6	89	61.3	25.8	5.4
7	93.5	75	54.3	35.9
8	91.3	72.8	41.3	6.5
9	76.1	64.1	50	41.3
10	48.9	48.9	48.9	48.9
11	90.2	82.6	33.7	16.3
12	22.8	33.7	9.8	9.7

Confidence level with Control				
Student #	1	3	5	7
1	22.8	3.2	28.3	0
2	98	97	95	90
3	87	92.4	88.6	78.3
4	80.4	95.6	72.8	81.5
5	51.1	61	63	43.5
6	89.1	86.9	70.6	67.4
7	76.1	67.4	51.1	24
8	91	62	42	41
9	51.1	73.9	42.4	12
10	95.7	87	77.1	88
11	92.4	67.4	44.6	35.9

Appendix C: SAS Code

The below mentioned SAS program is used to calculate the p-values for:

1. Results on Domain and Design Error

2. Percentage decrease in confidence for all domain populations

3. Effect of Controls on all Domain Populations

4. Confidence level comparison for all domain population with and without controls.

5. Pairwise comparisons between Domain Generalists & Non-Domain Generals,

   Domain Generalists & Domain Specialists

SAS code:

```
ODS RTF FILE'U:\TEST.RTF';
DATA NOCONTROLST;
INPUT Student $ T1 T3 T5 T7;
TYPE='STU';
Student =student+100;
DATALINES;
1          66.3      54.3      7.6       0
2          94.6      71.7      19.6      13
3          75        63.8      47.4      13.2
.          .         .         .         .
.          .         .         .         .
.          .         .         .         .
;
DATA T1;
SET NOCONTROLST;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET NOCONTROLST;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET NOCONTROLST;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET NOCONTROLST;
CONFIDENCE=T7;
```

```
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA NCST;
SET T1 T3 T5 T7;
CONTROL='NO ';
*PROC PRINT;
RUN;
QUIT;

DATA CONTROLST;
INPUT Student $ T1 T3 T5 T7;
TYPE='STU';
Student =student+200;
DATALINES;
1       22.8    13      12      7.6
2       96.7    76.1    70.7    50
3       97.8    80      65.2    51.1
;

DATA T1;
SET CONTROLST;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET CONTROLST;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET CONTROLST;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET CONTROLST;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA WCST;
SET T1 T3 T5 T7;
CONTROL='YES';

DATA NOCONTROLPSY;
INPUT Student $ T1 T3 T5 T7;
TYPE='PSY';
Student =student+300;
DATALINES;
1       63.4    39.8    18.3    9.7
2       78.5    61.3    24.7    6.5
3       45.2    32.3    16.1    8.6
;
DATA T1;
```

```
SET NOCONTROLPSY;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET NOCONTROLPSY;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET NOCONTROLPSY;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET NOCONTROLPSY;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA NCPSY;
SET T1 T3 T5 T7;
CONTROL='NO ';
*PROC PRINT;
RUN;
QUIT;

DATA CONTROLPSY;
INPUT Student $ T1 T3 T5 T7;
TYPE='PSY';
Student =student+400;
DATALINES;
1       87      79.6    77.4    76.3
2       52.7    40      20.4    9.6
3       74.2    65.6    23.7    5.4
.       .       .       .       .
.       .       .       .       .
.       .       .       .       .
;
DATA T1;
SET CONTROLPSY;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET CONTROLPSY;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET CONTROLPSY;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
```

```
DATA T7;
SET CONTROLPSY;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA WCPSY;
SET T1 T3 T5 T7;
CONTROL='YES';
DATA NOCONTROLEXP;
INPUT Student $ T1 T3 T5 T7;
TYPE='EXP';
Student =student+500;
DATALINES;
1       92.4    89.1    73.9    56.5
2       43.5    32.6    61      25
3       67.4    62      64      62
.       .       .       .       .
.       .       .       .       .
.       .       .       .       .
;
DATA T1;
SET NOCONTROLEXP;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET NOCONTROLEXP;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET NOCONTROLEXP;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET NOCONTROLEXP;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA NCEXP;
SET T1 T3 T5 T7;
CONTROL='NO ';
*PROC PRINT;
RUN;
QUIT;

DATA CONTROLEXP;
INPUT Student $ T1 T3 T5 T7;
TYPE='EXP';
Student =student+600;
DATALINES;
1       22.8    3.2     28.3    0
```

```
2         98        97        95        90
3         87        92.4      88.6      78.3
;
DATA T1;
SET CONTROLEXP;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET CONTROLEXP;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET CONTROLEXP;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET CONTROLEXP;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA WCEXP;
SET T1 T3 T5 T7;
CONTROL='YES';
DATA ALL;
SET NCST WCST ncpsy wcpsy ncexp wcexp;

PROC MIXED DATA=ALL;
CLASS TYPE CONTROL TASKS STUDENT;
MODEL CONFIDENCE=TASKS|CONTROL|TYPE / OUTP=RESIDUAL DDFM=SAT;
RANDOM STUDENT;

PROC MIXED DATA=ALL;
CLASS TYPE CONTROL TASKS STUDENT;
MODEL CONFIDENCE=TASKS|CONTROL|TYPE / OUTP=RESIDUAL DDFM=SAT;
RANDOM STUDENT;
REPEATED / TYPE=UN GROUP=TYPE*CONTROL;

PROC MIXED DATA=ALL;
CLASS TYPE CONTROL TASKS STUDENT;
MODEL CONFIDENCE=TASKS|CONTROL|TYPE / OUTP=RESIDUAL DDFM=SAT;
RANDOM STUDENT;
REPEATED / TYPE=UN GROUP=TYPE;
LSMEANS TASKS*CONTROL*TYPE / SLICE=TASKS*CONTROL SLICE=TASKS*TYPE pdiff;
ODS OUTPUT LSMEANS=LSMEAN;
/*
PROC UNIVARIATE NORMAL PLOT;
VAR RESID;
*/
GOPTIONS COLORS= (BLACK);
SYMBOL1 V=CIRCLE I=J L=1;
```

```
SYMBOL2 V=SQUARE I=J L=3;
SYMBOL3 V=DIAMOND I=J L=5;
PROC SORT DATA=LSMEAN; BY TYPE;
PROC GPLOT DATA=LSMEAN; BY TYPE;
PLOT ESTIMATE*TASKS=CONTROL;
PROC SORT DATA=LSMEAN; BY CONTROL;
PROC GPLOT DATA=LSMEAN; BY CONTROL;
PLOT ESTIMATE*TASKS=TYPE;

PROC MIXED DATA=ALL;
CLASS TYPE CONTROL TASKS STUDENT;
MODEL CONFIDENCE=TASKS|CONTROL|TYPE / OUTP=RESIDUAL DDFM=SAT;
RANDOM STUDENT;
REPEATED / TYPE=UN GROUP=TYPE;
LSMEANS TASKS*TYPE / SLICE=TASKS SLICE=TYPE pdiff;
LSMESTIMATE TASKS*TYPE 'EXP TASKS 1 VS 3' 1 -1;
LSMESTIMATE TASKS*TYPE 'EXP TASKS 3 VS 5' 0 1 -1;
LSMESTIMATE TASKS*TYPE 'EXP TASKS 5 VS 7' 0 0 1 -1;
LSMESTIMATE TASKS*TYPE 'PSY TASKS 1 VS 3' 0 0 0 0 1 -1;
LSMESTIMATE TASKS*TYPE 'PSY TASKS 3 VS 5' 0 0 0 0 0 1 -1;
LSMESTIMATE TASKS*TYPE 'PSY TASKS 5 VS 7' 0 0 0 0 0 0 1 -1;
LSMESTIMATE TASKS*TYPE 'STU TASKS 1 VS 3' 0 0 0 0 0 0 0 0 1 -1;
LSMESTIMATE TASKS*TYPE 'STU TASKS 3 VS 5' 0 0 0 0 0 0 0 0 0 1 -1;
LSMESTIMATE TASKS*TYPE 'STU TASKS 5 VS 7' 0 0 0 0 0 0 0 0 0 0 1 -1;
ODS OUTPUT LSMEANS=LSMEAN2;
PROC GPLOT DATA=LSMEAN2;
PLOT ESTIMATE*TASKS=TYPE;

PROC MIXED DATA=ALL;
CLASS TYPE CONTROL TASKS STUDENT;
MODEL CONFIDENCE=TASKS|CONTROL|TYPE / OUTP=RESIDUAL DDFM=SAT;
RANDOM STUDENT;
REPEATED / TYPE=UN GROUP=TYPE;
LSMEANS CONTROL*TASKS / SLICE=TASKS SLICE=CONTROL;
ODS OUTPUT LSMEANS=LSMEAN3;
PROC GPLOT DATA=LSMEAN3;
PLOT ESTIMATE*TASKS=CONTROL;
ODS RTF CLOSE;
RUN;
QUIT;
```

Model Information	
Data Set	WORK.ALL
Dependent Variable	CONFIDENCE
Covariance Structures	Variance Components, Unstructured
Group Effect	TYPE
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	**Levels**	**Values**
TYPE	3	EXP PSY STU
CONTROL	2	NO YES
TASKS	4	1 3 5 7

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
0	1	4629.42313897	
1	2	4365.89611566	0.00001803
2	1	4365.86472817	0.00000002
3	1	4365.86469279	0.00000000

Convergence criteria met.

Covariance Parameter Estimates		
Cov Parm	**Group**	**Estimate**
Student		217.47
UN(1,1)	TYPE EXP	351.49
UN(1,1)	TYPE PSY	191.19
UN(1,1)	TYPE STU	99.0268

Fit Statistics	
-2 Res Log Likelihood	4365.9
AIC (smaller is better)	4373.9
AICC (smaller is better)	4373.9
BIC (smaller is better)	4385.2

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
TASKS	3	139	137.88	<.0001
CONTROL	1	136	8.92	0.0033
CONTROL*TASKS	3	139	1.95	0.1240
TYPE	2	132	3.38	0.0369
TYPE*TASKS	6	98.1	4.24	0.0008
TYPE*CONTROL	2	132	0.52	0.5954
TYPE*CONTROL*TASKS	6	98.1	0.94	0.4683

Least Squares Means							
Effect	**TYPE**	**CONTROL**	**TASKS**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
TYPE*CONTROL*TASKS	EXP	NO	1	68.0665	7.8046	8.72	<.0001
TYPE*CONTROL*TASKS	EXP	NO	3	59.7415	7.8046	7.65	<.0001
TYPE*CONTROL*TASKS	EXP	NO	5	44.3998	7.8046	5.69	<.0001
TYPE*CONTROL*TASKS	EXP	NO	7	30.2581	7.8046	3.88	0.0002
TYPE*CONTROL*TASKS	EXP	YES	1	77.6584	7.9852	9.73	<.0001
TYPE*CONTROL*TASKS	EXP	YES	3	73.9402	7.9852	9.26	<.0001
TYPE*CONTROL*TASKS	EXP	YES	5	63.1856	7.9852	7.91	<.0001
TYPE*CONTROL*TASKS	EXP	YES	7	52.8311	7.9852	6.62	<.0001
TYPE*CONTROL*TASKS	PSY	NO	1	67.9600	5.2195	13.02	<.0001
TYPE*CONTROL*TASKS	PSY	NO	3	44.7133	5.2195	8.57	<.0001
TYPE*CONTROL*TASKS	PSY	NO	5	32.9733	5.2195	6.32	<.0001
TYPE*CONTROL*TASKS	PSY	NO	7	21.4467	5.2195	4.11	<.0001
TYPE*CONTROL*TASKS	PSY	YES	1	70.3000	5.4027	13.01	<.0001
TYPE*CONTROL*TASKS	PSY	YES	3	65.3743	5.4027	12.10	<.0001
TYPE*CONTROL*TASKS	PSY	YES	5	43.7793	5.4027	8.10	<.0001
TYPE*CONTROL*TASKS	PSY	YES	7	28.9364	5.4027	5.36	<.0001
TYPE*CONTROL*TASKS	STU	NO	1	74.4550	2.8129	26.47	<.0001
TYPE*CONTROL*TASKS	STU	NO	3	50.1625	2.8129	17.83	<.0001
TYPE*CONTROL*TASKS	STU	NO	5	31.9525	2.8129	11.36	<.0001
TYPE*CONTROL*TASKS	STU	NO	7	15.3475	2.8129	5.46	<.0001
TYPE*CONTROL*TASKS	STU	YES	1	76.0636	2.6820	28.36	<.0001
TYPE*CONTROL*TASKS	STU	YES	3	56.7341	2.6820	21.15	<.0001
TYPE*CONTROL*TASKS	STU	YES	5	42.2295	2.6820	15.75	<.0001
TYPE*CONTROL*TASKS	STU	YES	7	25.0523	2.6820	9.34	<.0001

Differences of Least Squares Means								
TYPE	CONTROL	TASKS	TYPE	CONTROL	TASKS	Estimate	Standard Error	DF
EXP	NO	1	EXP	NO	3	8.3250	7.6539	71.4
EXP	NO	1	EXP	NO	5	23.6667	7.6539	71.4
EXP	NO	1	EXP	NO	7	37.8083	7.6539	71.4
EXP	NO	1	EXP	YES	1	-9.5919	11.1658	166
EXP	NO	1	EXP	YES	3	-5.8737	11.1658	166
EXP	NO	1	EXP	YES	5	4.8808	11.1658	166
EXP	NO	1	EXP	YES	7	15.2354	11.1658	166
EXP	NO	1	PSY	NO	1	0.1065	9.3891	219
EXP	NO	1	PSY	NO	3	23.3531	9.3891	219
EXP	NO	1	PSY	NO	5	35.0931	9.3891	219
EXP	NO	1	PSY	NO	7	46.6198	9.3891	219
EXP	NO	1	PSY	YES	1	-2.2335	9.4922	221
EXP	NO	1	PSY	YES	3	2.6922	9.4922	221
EXP	NO	1	PSY	YES	5	24.2872	9.4922	221
EXP	NO	1	PSY	YES	7	39.1300	9.4922	221
EXP	NO	1	STU	NO	1	-6.3885	8.2961	181
EXP	NO	1	STU	NO	3	17.9040	8.2961	181
EXP	NO	1	STU	NO	5	36.1140	8.2961	181
EXP	NO	1	STU	NO	7	52.7190	8.2961	181
EXP	NO	1	STU	YES	1	-7.9972	8.2526	180
EXP	NO	1	STU	YES	3	11.3324	8.2526	180
EXP	NO	1	STU	YES	5	25.8369	8.2526	180
EXP	NO	1	STU	YES	7	43.0142	8.2526	180
EXP	NO	3	EXP	NO	5	15.3417	7.6539	71.4
EXP	NO	3	EXP	NO	7	29.4833	7.6539	71.4
EXP	NO	3	EXP	YES	1	-17.9169	11.1658	166
EXP	NO	3	EXP	YES	3	-14.1987	11.1658	166
EXP	NO	3	EXP	YES	5	-3.4442	11.1658	166
EXP	NO	3	EXP	YES	7	6.9104	11.1658	166
EXP	NO	3	PSY	NO	1	-8.2185	9.3891	219
EXP	NO	3	PSY	NO	3	15.0281	9.3891	219

Differences of Least Squares Means

TYPE	CONTROL	TASKS	TYPE	CONTROL	TASKS	Estimate	Standard Error	DF
EXP	NO	3	PSY	NO	5	26.7681	9.3891	219
EXP	NO	3	PSY	NO	7	38.2948	9.3891	219
EXP	NO	3	PSY	YES	1	-10.5585	9.4922	221
EXP	NO	3	PSY	YES	3	-5.6328	9.4922	221
EXP	NO	3	PSY	YES	5	15.9622	9.4922	221
EXP	NO	3	PSY	YES	7	30.8050	9.4922	221
EXP	NO	3	STU	NO	1	-14.7135	8.2961	181
EXP	NO	3	STU	NO	3	9.5790	8.2961	181
EXP	NO	3	STU	NO	5	27.7890	8.2961	181
EXP	NO	3	STU	NO	7	44.3940	8.2961	181
EXP	NO	3	STU	YES	1	-16.3222	8.2526	180
EXP	NO	3	STU	YES	3	3.0074	8.2526	180
EXP	NO	3	STU	YES	5	17.5119	8.2526	180
EXP	NO	3	STU	YES	7	34.6892	8.2526	180
EXP	NO	5	EXP	NO	7	14.1417	7.6539	71.4
EXP	NO	5	EXP	YES	1	-33.2586	11.1658	166
EXP	NO	5	EXP	YES	3	-29.5404	11.1658	166
EXP	NO	5	EXP	YES	5	-18.7858	11.1658	166
EXP	NO	5	EXP	YES	7	-8.4313	11.1658	166
EXP	NO	5	PSY	NO	1	-23.5602	9.3891	219
EXP	NO	5	PSY	NO	3	-0.3135	9.3891	219
EXP	NO	5	PSY	NO	5	11.4265	9.3891	219
EXP	NO	5	PSY	NO	7	22.9531	9.3891	219
EXP	NO	5	PSY	YES	1	-25.9002	9.4922	221
EXP	NO	5	PSY	YES	3	-20.9745	9.4922	221
EXP	NO	5	PSY	YES	5	0.6205	9.4922	221
EXP	NO	5	PSY	YES	7	15.4634	9.4922	221
EXP	NO	5	STU	NO	1	-30.0552	8.2961	181
EXP	NO	5	STU	NO	3	-5.7627	8.2961	181
EXP	NO	5	STU	NO	5	12.4473	8.2961	181
EXP	NO	5	STU	NO	7	29.0523	8.2961	181

Differences of Least Squares Means								
TYPE	CONTROL	TASKS	TYPE	CONTROL	TASKS	Estimate	Standard Error	DF
EXP	NO	5	STU	YES	1	-31.6638	8.2526	180
EXP	NO	5	STU	YES	3	-12.3343	8.2526	180
EXP	NO	5	STU	YES	5	2.1703	8.2526	180
EXP	NO	5	STU	YES	7	19.3475	8.2526	180
EXP	NO	7	EXP	YES	1	-47.4002	11.1658	166
EXP	NO	7	EXP	YES	3	-43.6820	11.1658	166
EXP	NO	7	EXP	YES	5	-32.9275	11.1658	166
EXP	NO	7	EXP	YES	7	-22.5730	11.1658	166
EXP	NO	7	PSY	NO	1	-37.7019	9.3891	219
EXP	NO	7	PSY	NO	3	-14.4552	9.3891	219
EXP	NO	7	PSY	NO	5	-2.7152	9.3891	219
EXP	NO	7	PSY	NO	7	8.8115	9.3891	219
EXP	NO	7	PSY	YES	1	-40.0419	9.4922	221
EXP	NO	7	PSY	YES	3	-35.1161	9.4922	221
EXP	NO	7	PSY	YES	5	-13.5211	9.4922	221
EXP	NO	7	PSY	YES	7	1.3217	9.4922	221
EXP	NO	7	STU	NO	1	-44.1969	8.2961	181
EXP	NO	7	STU	NO	3	-19.9044	8.2961	181
EXP	NO	7	STU	NO	5	-1.6944	8.2961	181
EXP	NO	7	STU	NO	7	14.9106	8.2961	181
EXP	NO	7	STU	YES	1	-45.8055	8.2526	180
EXP	NO	7	STU	YES	3	-26.4759	8.2526	180
EXP	NO	7	STU	YES	5	-11.9714	8.2526	180
EXP	NO	7	STU	YES	7	5.2059	8.2526	180
EXP	YES	1	EXP	YES	3	3.7182	7.9942	71.4
EXP	YES	1	EXP	YES	5	14.4727	7.9942	71.4
EXP	YES	1	EXP	YES	7	24.8273	7.9942	71.4
EXP	YES	1	PSY	NO	1	9.6984	9.5397	218
EXP	YES	1	PSY	NO	3	32.9450	9.5397	218
EXP	YES	1	PSY	NO	5	44.6850	9.5397	218
EXP	YES	1	PSY	NO	7	56.2117	9.5397	218

Differences of Least Squares Means								
TYPE	CONTROL	TASKS	TYPE	CONTROL	TASKS	Estimate	Standard Error	DF
EXP	YES	1	PSY	YES	1	7.3584	9.6412	221
EXP	YES	1	PSY	YES	3	12.2841	9.6412	221
EXP	YES	1	PSY	YES	5	33.8791	9.6412	221
EXP	YES	1	PSY	YES	7	48.7219	9.6412	221
EXP	YES	1	STU	NO	1	3.2034	8.4661	180
EXP	YES	1	STU	NO	3	27.4959	8.4661	180
EXP	YES	1	STU	NO	5	45.7059	8.4661	180
EXP	YES	1	STU	NO	7	62.3109	8.4661	180
EXP	YES	1	STU	YES	1	1.5947	8.4235	179
EXP	YES	1	STU	YES	3	20.9243	8.4235	179
EXP	YES	1	STU	YES	5	35.4288	8.4235	179
EXP	YES	1	STU	YES	7	52.6061	8.4235	179
EXP	YES	3	EXP	YES	5	10.7545	7.9942	71.4
EXP	YES	3	EXP	YES	7	21.1091	7.9942	71.4
EXP	YES	3	PSY	NO	1	5.9802	9.5397	218
EXP	YES	3	PSY	NO	3	29.2269	9.5397	218
EXP	YES	3	PSY	NO	5	40.9669	9.5397	218
EXP	YES	3	PSY	NO	7	52.4935	9.5397	218
EXP	YES	3	PSY	YES	1	3.6402	9.6412	221
EXP	YES	3	PSY	YES	3	8.5659	9.6412	221
EXP	YES	3	PSY	YES	5	30.1609	9.6412	221
EXP	YES	3	PSY	YES	7	45.0038	9.6412	221
EXP	YES	3	STU	NO	1	-0.5148	8.4661	180
EXP	YES	3	STU	NO	3	23.7777	8.4661	180
EXP	YES	3	STU	NO	5	41.9877	8.4661	180
EXP	YES	3	STU	NO	7	58.5927	8.4661	180
EXP	YES	3	STU	YES	1	-2.1234	8.4235	179
EXP	YES	3	STU	YES	3	17.2061	8.4235	179
EXP	YES	3	STU	YES	5	31.7106	8.4235	179
EXP	YES	3	STU	YES	7	48.8879	8.4235	179
EXP	YES	5	EXP	YES	7	10.3545	7.9942	71.4

Differences of Least Squares Means

TYPE	CONTROL	TASKS	TYPE	CONTROL	TASKS	Estimate	Standard Error	DF
EXP	YES	5	PSY	NO	1	-4.7744	9.5397	218
EXP	YES	5	PSY	NO	3	18.4723	9.5397	218
EXP	YES	5	PSY	NO	5	30.2123	9.5397	218
EXP	YES	5	PSY	NO	7	41.7390	9.5397	218
EXP	YES	5	PSY	YES	1	-7.1144	9.6412	221
EXP	YES	5	PSY	YES	3	-2.1886	9.6412	221
EXP	YES	5	PSY	YES	5	19.4064	9.6412	221
EXP	YES	5	PSY	YES	7	34.2492	9.6412	221
EXP	YES	5	STU	NO	1	-11.2694	8.4661	180
EXP	YES	5	STU	NO	3	13.0231	8.4661	180
EXP	YES	5	STU	NO	5	31.2331	8.4661	180
EXP	YES	5	STU	NO	7	47.8381	8.4661	180
EXP	YES	5	STU	YES	1	-12.8780	8.4235	179
EXP	YES	5	STU	YES	3	6.4516	8.4235	179
EXP	YES	5	STU	YES	5	20.9561	8.4235	179
EXP	YES	5	STU	YES	7	38.1334	8.4235	179
EXP	YES	7	PSY	NO	1	-15.1289	9.5397	218
EXP	YES	7	PSY	NO	3	8.1178	9.5397	218
EXP	YES	7	PSY	NO	5	19.8578	9.5397	218
EXP	YES	7	PSY	NO	7	31.3844	9.5397	218
EXP	YES	7	PSY	YES	1	-17.4689	9.6412	221
EXP	YES	7	PSY	YES	3	-12.5432	9.6412	221
EXP	YES	7	PSY	YES	5	9.0518	9.6412	221
EXP	YES	7	PSY	YES	7	23.8947	9.6412	221
EXP	YES	7	STU	NO	1	-21.6239	8.4661	180
EXP	YES	7	STU	NO	3	2.6686	8.4661	180
EXP	YES	7	STU	NO	5	20.8786	8.4661	180
EXP	YES	7	STU	NO	7	37.4836	8.4661	180
EXP	YES	7	STU	YES	1	-23.2325	8.4235	179
EXP	YES	7	STU	YES	3	-3.9030	8.4235	179
EXP	YES	7	STU	YES	5	10.6016	8.4235	179

Differences of Least Squares Means								
TYPE	CONTROL	TASKS	TYPE	CONTROL	TASKS	Estimate	Standard Error	DF
EXP	YES	7	STU	YES	7	27.7788	8.4235	179
PSY	NO	1	PSY	NO	3	23.2467	5.0489	81.8
PSY	NO	1	PSY	NO	5	34.9867	5.0489	81.8
PSY	NO	1	PSY	NO	7	46.5133	5.0489	81.8
PSY	NO	1	PSY	YES	1	-2.3400	7.5122	181
PSY	NO	1	PSY	YES	3	2.5857	7.5122	181
PSY	NO	1	PSY	YES	5	24.1807	7.5122	181
PSY	NO	1	PSY	YES	7	39.0236	7.5122	181
PSY	NO	1	STU	NO	1	-6.4950	5.9292	203
PSY	NO	1	STU	NO	3	17.7975	5.9292	203
PSY	NO	1	STU	NO	5	36.0075	5.9292	203
PSY	NO	1	STU	NO	7	52.6125	5.9292	203
PSY	NO	1	STU	YES	1	-8.1036	5.8683	201
PSY	NO	1	STU	YES	3	11.2259	5.8683	201
PSY	NO	1	STU	YES	5	25.7305	5.8683	201
PSY	NO	1	STU	YES	7	42.9077	5.8683	201
PSY	NO	3	PSY	NO	5	11.7400	5.0489	81.8
PSY	NO	3	PSY	NO	7	23.2667	5.0489	81.8
PSY	NO	3	PSY	YES	1	-25.5867	7.5122	181
PSY	NO	3	PSY	YES	3	-20.6610	7.5122	181
PSY	NO	3	PSY	YES	5	0.9340	7.5122	181
PSY	NO	3	PSY	YES	7	15.7769	7.5122	181
PSY	NO	3	STU	NO	1	-29.7417	5.9292	203
PSY	NO	3	STU	NO	3	-5.4492	5.9292	203
PSY	NO	3	STU	NO	5	12.7608	5.9292	203
PSY	NO	3	STU	NO	7	29.3658	5.9292	203
PSY	NO	3	STU	YES	1	-31.3503	5.8683	201
PSY	NO	3	STU	YES	3	-12.0208	5.8683	201
PSY	NO	3	STU	YES	5	2.4838	5.8683	201
PSY	NO	3	STU	YES	7	19.6611	5.8683	201
PSY	NO	5	PSY	NO	7	11.5267	5.0489	81.8

Differences of Least Squares Means								
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**TASKS**	**Estimate**	**Standard Error**	**DF**
PSY	NO	5	PSY	YES	1	-37.3267	7.5122	181
PSY	NO	5	PSY	YES	3	-32.4010	7.5122	181
PSY	NO	5	PSY	YES	5	-10.8060	7.5122	181
PSY	NO	5	PSY	YES	7	4.0369	7.5122	181
PSY	NO	5	STU	NO	1	-41.4817	5.9292	203
PSY	NO	5	STU	NO	3	-17.1892	5.9292	203
PSY	NO	5	STU	NO	5	1.0208	5.9292	203
PSY	NO	5	STU	NO	7	17.6258	5.9292	203
PSY	NO	5	STU	YES	1	-43.0903	5.8683	201
PSY	NO	5	STU	YES	3	-23.7608	5.8683	201
PSY	NO	5	STU	YES	5	-9.2562	5.8683	201
PSY	NO	5	STU	YES	7	7.9211	5.8683	201
PSY	NO	7	PSY	YES	1	-48.8533	7.5122	181
PSY	NO	7	PSY	YES	3	-43.9276	7.5122	181
PSY	NO	7	PSY	YES	5	-22.3326	7.5122	181
PSY	NO	7	PSY	YES	7	-7.4898	7.5122	181
PSY	NO	7	STU	NO	1	-53.0083	5.9292	203
PSY	NO	7	STU	NO	3	-28.7158	5.9292	203
PSY	NO	7	STU	NO	5	-10.5058	5.9292	203
PSY	NO	7	STU	NO	7	6.0992	5.9292	203
PSY	NO	7	STU	YES	1	-54.6170	5.8683	201
PSY	NO	7	STU	YES	3	-35.2874	5.8683	201
PSY	NO	7	STU	YES	5	-20.7829	5.8683	201
PSY	NO	7	STU	YES	7	-3.6056	5.8683	201
PSY	YES	1	PSY	YES	3	4.9257	5.2261	81.8
PSY	YES	1	PSY	YES	5	26.5207	5.2261	81.8
PSY	YES	1	PSY	YES	7	41.3636	5.2261	81.8
PSY	YES	1	STU	NO	1	-4.1550	6.0911	202
PSY	YES	1	STU	NO	3	20.1375	6.0911	202
PSY	YES	1	STU	NO	5	38.3475	6.0911	202
PSY	YES	1	STU	NO	7	54.9525	6.0911	202

Differences of Least Squares Means								
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**TASKS**	**Estimate**	**Standard Error**	**DF**
PSY	YES	1	STU	YES	1	-5.7636	6.0318	200
PSY	YES	1	STU	YES	3	13.5659	6.0318	200
PSY	YES	1	STU	YES	5	28.0705	6.0318	200
PSY	YES	1	STU	YES	7	45.2477	6.0318	200
PSY	YES	3	PSY	YES	5	21.5950	5.2261	81.8
PSY	YES	3	PSY	YES	7	36.4379	5.2261	81.8
PSY	YES	3	STU	NO	1	-9.0807	6.0911	202
PSY	YES	3	STU	NO	3	15.2118	6.0911	202
PSY	YES	3	STU	NO	5	33.4218	6.0911	202
PSY	YES	3	STU	NO	7	50.0268	6.0911	202
PSY	YES	3	STU	YES	1	-10.6894	6.0318	200
PSY	YES	3	STU	YES	3	8.6402	6.0318	200
PSY	YES	3	STU	YES	5	23.1447	6.0318	200
PSY	YES	3	STU	YES	7	40.3220	6.0318	200
PSY	YES	5	PSY	YES	7	14.8429	5.2261	81.8
PSY	YES	5	STU	NO	1	-30.6757	6.0911	202
PSY	YES	5	STU	NO	3	-6.3832	6.0911	202
PSY	YES	5	STU	NO	5	11.8268	6.0911	202
PSY	YES	5	STU	NO	7	28.4318	6.0911	202
PSY	YES	5	STU	YES	1	-32.2844	6.0318	200
PSY	YES	5	STU	YES	3	-12.9548	6.0318	200
PSY	YES	5	STU	YES	5	1.5497	6.0318	200
PSY	YES	5	STU	YES	7	18.7270	6.0318	200
PSY	YES	7	STU	NO	1	-45.5186	6.0911	202
PSY	YES	7	STU	NO	3	-21.2261	6.0911	202
PSY	YES	7	STU	NO	5	-3.0161	6.0911	202
PSY	YES	7	STU	NO	7	13.5889	6.0911	202
PSY	YES	7	STU	YES	1	-47.1272	6.0318	200
PSY	YES	7	STU	YES	3	-27.7977	6.0318	200
PSY	YES	7	STU	YES	5	-13.2931	6.0318	200
PSY	YES	7	STU	YES	7	3.8842	6.0318	200

Differences of Least Squares Means								
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**TASKS**	**Estimate**	**Standard Error**	**DF**
STU	NO	1	STU	NO	3	24.2925	2.2252	247
STU	NO	1	STU	NO	5	42.5025	2.2252	247
STU	NO	1	STU	NO	7	59.1075	2.2252	247
STU	NO	1	STU	YES	1	-1.6086	3.8866	183
STU	NO	1	STU	YES	3	17.7209	3.8866	183
STU	NO	1	STU	YES	5	32.2255	3.8866	183
STU	NO	1	STU	YES	7	49.4027	3.8866	183
STU	NO	3	STU	NO	5	18.2100	2.2252	247
STU	NO	3	STU	NO	7	34.8150	2.2252	247
STU	NO	3	STU	YES	1	-25.9011	3.8866	183
STU	NO	3	STU	YES	3	-6.5716	3.8866	183
STU	NO	3	STU	YES	5	7.9330	3.8866	183
STU	NO	3	STU	YES	7	25.1102	3.8866	183
STU	NO	5	STU	NO	7	16.6050	2.2252	247
STU	NO	5	STU	YES	1	-44.1111	3.8866	183
STU	NO	5	STU	YES	3	-24.7816	3.8866	183
STU	NO	5	STU	YES	5	-10.2770	3.8866	183
STU	NO	5	STU	YES	7	6.9002	3.8866	183
STU	NO	7	STU	YES	1	-60.7161	3.8866	183
STU	NO	7	STU	YES	3	-41.3866	3.8866	183
STU	NO	7	STU	YES	5	-26.8820	3.8866	183
STU	NO	7	STU	YES	7	-9.7048	3.8866	183
STU	YES	1	STU	YES	3	19.3295	2.1216	247
STU	YES	1	STU	YES	5	33.8341	2.1216	247
STU	YES	1	STU	YES	7	51.0114	2.1216	247
STU	YES	3	STU	YES	5	14.5045	2.1216	247
STU	YES	3	STU	YES	7	31.6818	2.1216	247
STU	YES	5	STU	YES	7	17.1773	2.1216	247

Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
EXP	NO	1	EXP	NO	1.09	0.2804
EXP	NO	1	EXP	NO	3.09	0.0028
EXP	NO	1	EXP	NO	4.94	<.0001
EXP	NO	1	EXP	YES	-0.86	0.3916
EXP	NO	1	EXP	YES	-0.53	0.5996
EXP	NO	1	EXP	YES	0.44	0.6626
EXP	NO	1	EXP	YES	1.36	0.1743
EXP	NO	1	PSY	NO	0.01	0.9910
EXP	NO	1	PSY	NO	2.49	0.0136
EXP	NO	1	PSY	NO	3.74	0.0002
EXP	NO	1	PSY	NO	4.97	<.0001
EXP	NO	1	PSY	YES	-0.24	0.8142
EXP	NO	1	PSY	YES	0.28	0.7770
EXP	NO	1	PSY	YES	2.56	0.0112
EXP	NO	1	PSY	YES	4.12	<.0001
EXP	NO	1	STU	NO	-0.77	0.4423
EXP	NO	1	STU	NO	2.16	0.0322
EXP	NO	1	STU	NO	4.35	<.0001
EXP	NO	1	STU	NO	6.35	<.0001
EXP	NO	1	STU	YES	-0.97	0.3338
EXP	NO	1	STU	YES	1.37	0.1714
EXP	NO	1	STU	YES	3.13	0.0020
EXP	NO	1	STU	YES	5.21	<.0001
EXP	NO	3	EXP	NO	2.00	0.0488
EXP	NO	3	EXP	NO	3.85	0.0003
EXP	NO	3	EXP	YES	-1.60	0.1105
EXP	NO	3	EXP	YES	-1.27	0.2053
EXP	NO	3	EXP	YES	-0.31	0.7581
EXP	NO	3	EXP	YES	0.62	0.5368
EXP	NO	3	PSY	NO	-0.88	0.3824
EXP	NO	3	PSY	NO	1.60	0.1109
EXP	NO	3	PSY	NO	2.85	0.0048

Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
EXP	NO	3	PSY	NO	4.08	<.0001
EXP	NO	3	PSY	YES	-1.11	0.2672
EXP	NO	3	PSY	YES	-0.59	0.5535
EXP	NO	3	PSY	YES	1.68	0.0941
EXP	NO	3	PSY	YES	3.25	0.0014
EXP	NO	3	STU	NO	-1.77	0.0778
EXP	NO	3	STU	NO	1.15	0.2498
EXP	NO	3	STU	NO	3.35	0.0010
EXP	NO	3	STU	NO	5.35	<.0001
EXP	NO	3	STU	YES	-1.98	0.0495
EXP	NO	3	STU	YES	0.36	0.7160
EXP	NO	3	STU	YES	2.12	0.0352
EXP	NO	3	STU	YES	4.20	<.0001
EXP	NO	5	EXP	NO	1.85	0.0688
EXP	NO	5	EXP	YES	-2.98	0.0033
EXP	NO	5	EXP	YES	-2.65	0.0089
EXP	NO	5	EXP	YES	-1.68	0.0944
EXP	NO	5	EXP	YES	-0.76	0.4513
EXP	NO	5	PSY	NO	-2.51	0.0128
EXP	NO	5	PSY	NO	-0.03	0.9734
EXP	NO	5	PSY	NO	1.22	0.2249
EXP	NO	5	PSY	NO	2.44	0.0153
EXP	NO	5	PSY	YES	-2.73	0.0069
EXP	NO	5	PSY	YES	-2.21	0.0282
EXP	NO	5	PSY	YES	0.07	0.9479
EXP	NO	5	PSY	YES	1.63	0.1047
EXP	NO	5	STU	NO	-3.62	0.0004
EXP	NO	5	STU	NO	-0.69	0.4882
EXP	NO	5	STU	NO	1.50	0.1353
EXP	NO	5	STU	NO	3.50	0.0006
EXP	NO	5	STU	YES	-3.84	0.0002
EXP	NO	5	STU	YES	-1.49	0.1368

Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
EXP	NO	5	STU	YES	0.26	0.7929
EXP	NO	5	STU	YES	2.34	0.0201
EXP	NO	7	EXP	YES	-4.25	<.0001
EXP	NO	7	EXP	YES	-3.91	0.0001
EXP	NO	7	EXP	YES	-2.95	0.0036
EXP	NO	7	EXP	YES	-2.02	0.0448
EXP	NO	7	PSY	NO	-4.02	<.0001
EXP	NO	7	PSY	NO	-1.54	0.1251
EXP	NO	7	PSY	NO	-0.29	0.7727
EXP	NO	7	PSY	NO	0.94	0.3490
EXP	NO	7	PSY	YES	-4.22	<.0001
EXP	NO	7	PSY	YES	-3.70	0.0003
EXP	NO	7	PSY	YES	-1.42	0.1557
EXP	NO	7	PSY	YES	0.14	0.8894
EXP	NO	7	STU	NO	-5.33	<.0001
EXP	NO	7	STU	NO	-2.40	0.0174
EXP	NO	7	STU	NO	-0.20	0.8384
EXP	NO	7	STU	NO	1.80	0.0740
EXP	NO	7	STU	YES	-5.55	<.0001
EXP	NO	7	STU	YES	-3.21	0.0016
EXP	NO	7	STU	YES	-1.45	0.1486
EXP	NO	7	STU	YES	0.63	0.5290
EXP	YES	1	EXP	YES	0.47	0.6433
EXP	YES	1	EXP	YES	1.81	0.0744
EXP	YES	1	EXP	YES	3.11	0.0027
EXP	YES	1	PSY	NO	1.02	0.3105
EXP	YES	1	PSY	NO	3.45	0.0007
EXP	YES	1	PSY	NO	4.68	<.0001
EXP	YES	1	PSY	NO	5.89	<.0001
EXP	YES	1	PSY	YES	0.76	0.4461
EXP	YES	1	PSY	YES	1.27	0.2040
EXP	YES	1	PSY	YES	3.51	0.0005

Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
EXP	YES	1	PSY	YES	5.05	<.0001
EXP	YES	1	STU	NO	0.38	0.7056
EXP	YES	1	STU	NO	3.25	0.0014
EXP	YES	1	STU	NO	5.40	<.0001
EXP	YES	1	STU	NO	7.36	<.0001
EXP	YES	1	STU	YES	0.19	0.8501
EXP	YES	1	STU	YES	2.48	0.0139
EXP	YES	1	STU	YES	4.21	<.0001
EXP	YES	1	STU	YES	6.25	<.0001
EXP	YES	3	EXP	YES	1.35	0.1828
EXP	YES	3	EXP	YES	2.64	0.0102
EXP	YES	3	PSY	NO	0.63	0.5314
EXP	YES	3	PSY	NO	3.06	0.0025
EXP	YES	3	PSY	NO	4.29	<.0001
EXP	YES	3	PSY	NO	5.50	<.0001
EXP	YES	3	PSY	YES	0.38	0.7061
EXP	YES	3	PSY	YES	0.89	0.3753
EXP	YES	3	PSY	YES	3.13	0.0020
EXP	YES	3	PSY	YES	4.67	<.0001
EXP	YES	3	STU	NO	-0.06	0.9516
EXP	YES	3	STU	NO	2.81	0.0055
EXP	YES	3	STU	NO	4.96	<.0001
EXP	YES	3	STU	NO	6.92	<.0001
EXP	YES	3	STU	YES	-0.25	0.8013
EXP	YES	3	STU	YES	2.04	0.0426
EXP	YES	3	STU	YES	3.76	0.0002
EXP	YES	3	STU	YES	5.80	<.0001
EXP	YES	5	EXP	YES	1.30	0.1994
EXP	YES	5	PSY	NO	-0.50	0.6172
EXP	YES	5	PSY	NO	1.94	0.0541
EXP	YES	5	PSY	NO	3.17	0.0018
EXP	YES	5	PSY	NO	4.38	<.0001

Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
EXP	YES	5	PSY	YES	-0.74	0.4614
EXP	YES	5	PSY	YES	-0.23	0.8206
EXP	YES	5	PSY	YES	2.01	0.0453
EXP	YES	5	PSY	YES	3.55	0.0005
EXP	YES	5	STU	NO	-1.33	0.1848
EXP	YES	5	STU	NO	1.54	0.1257
EXP	YES	5	STU	NO	3.69	0.0003
EXP	YES	5	STU	NO	5.65	<.0001
EXP	YES	5	STU	YES	-1.53	0.1281
EXP	YES	5	STU	YES	0.77	0.4447
EXP	YES	5	STU	YES	2.49	0.0138
EXP	YES	5	STU	YES	4.53	<.0001
EXP	YES	7	PSY	NO	-1.59	0.1142
EXP	YES	7	PSY	NO	0.85	0.3957
EXP	YES	7	PSY	NO	2.08	0.0385
EXP	YES	7	PSY	NO	3.29	0.0012
EXP	YES	7	PSY	YES	-1.81	0.0714
EXP	YES	7	PSY	YES	-1.30	0.1946
EXP	YES	7	PSY	YES	0.94	0.3488
EXP	YES	7	PSY	YES	2.48	0.0139
EXP	YES	7	STU	NO	-2.55	0.0115
EXP	YES	7	STU	NO	0.32	0.7530
EXP	YES	7	STU	NO	2.47	0.0146
EXP	YES	7	STU	NO	4.43	<.0001
EXP	YES	7	STU	YES	-2.76	0.0064
EXP	YES	7	STU	YES	-0.46	0.6437
EXP	YES	7	STU	YES	1.26	0.2098
EXP	YES	7	STU	YES	3.30	0.0012
PSY	NO	1	PSY	NO	4.60	<.0001
PSY	NO	1	PSY	NO	6.93	<.0001
PSY	NO	1	PSY	NO	9.21	<.0001
PSY	NO	1	PSY	YES	-0.31	0.7558

Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
PSY	NO	1	PSY	YES	0.34	0.7311
PSY	NO	1	PSY	YES	3.22	0.0015
PSY	NO	1	PSY	YES	5.19	<.0001
PSY	NO	1	STU	NO	-1.10	0.2746
PSY	NO	1	STU	NO	3.00	0.0030
PSY	NO	1	STU	NO	6.07	<.0001
PSY	NO	1	STU	NO	8.87	<.0001
PSY	NO	1	STU	YES	-1.38	0.1688
PSY	NO	1	STU	YES	1.91	0.0572
PSY	NO	1	STU	YES	4.38	<.0001
PSY	NO	1	STU	YES	7.31	<.0001
PSY	NO	3	PSY	NO	2.33	0.0225
PSY	NO	3	PSY	NO	4.61	<.0001
PSY	NO	3	PSY	YES	-3.41	0.0008
PSY	NO	3	PSY	YES	-2.75	0.0066
PSY	NO	3	PSY	YES	0.12	0.9012
PSY	NO	3	PSY	YES	2.10	0.0371
PSY	NO	3	STU	NO	-5.02	<.0001
PSY	NO	3	STU	NO	-0.92	0.3592
PSY	NO	3	STU	NO	2.15	0.0326
PSY	NO	3	STU	NO	4.95	<.0001
PSY	NO	3	STU	YES	-5.34	<.0001
PSY	NO	3	STU	YES	-2.05	0.0418
PSY	NO	3	STU	YES	0.42	0.6726
PSY	NO	3	STU	YES	3.35	0.0010
PSY	NO	5	PSY	NO	2.28	0.0250
PSY	NO	5	PSY	YES	-4.97	<.0001
PSY	NO	5	PSY	YES	-4.31	<.0001
PSY	NO	5	PSY	YES	-1.44	0.1520
PSY	NO	5	PSY	YES	0.54	0.5917
PSY	NO	5	STU	NO	-7.00	<.0001
PSY	NO	5	STU	NO	-2.90	0.0042

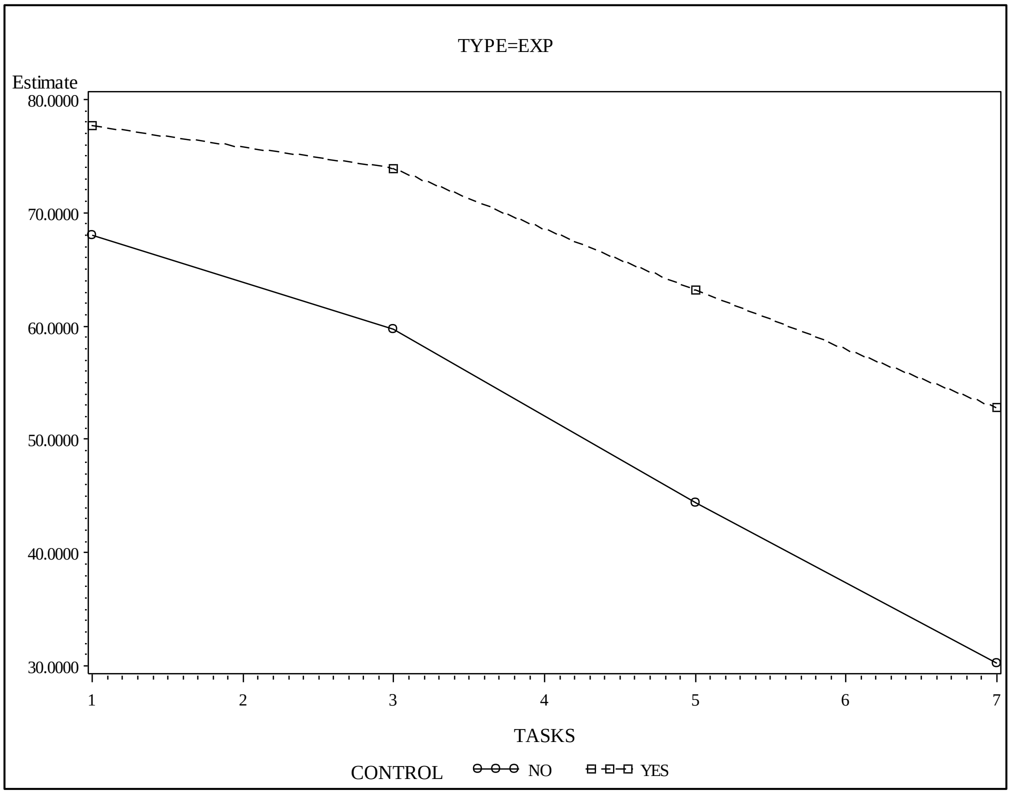
Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
PSY	NO	5	STU	NO	0.17	0.8635
PSY	NO	5	STU	NO	2.97	0.0033
PSY	NO	5	STU	YES	-7.34	<.0001
PSY	NO	5	STU	YES	-4.05	<.0001
PSY	NO	5	STU	YES	-1.58	0.1163
PSY	NO	5	STU	YES	1.35	0.1786
PSY	NO	7	PSY	YES	-6.50	<.0001
PSY	NO	7	PSY	YES	-5.85	<.0001
PSY	NO	7	PSY	YES	-2.97	0.0034
PSY	NO	7	PSY	YES	-1.00	0.3201
PSY	NO	7	STU	NO	-8.94	<.0001
PSY	NO	7	STU	NO	-4.84	<.0001
PSY	NO	7	STU	NO	-1.77	0.0779
PSY	NO	7	STU	NO	1.03	0.3049
PSY	NO	7	STU	YES	-9.31	<.0001
PSY	NO	7	STU	YES	-6.01	<.0001
PSY	NO	7	STU	YES	-3.54	0.0005
PSY	NO	7	STU	YES	-0.61	0.5396
PSY	YES	1	PSY	YES	0.94	0.3487
PSY	YES	1	PSY	YES	5.07	<.0001
PSY	YES	1	PSY	YES	7.91	<.0001
PSY	YES	1	STU	NO	-0.68	0.4959
PSY	YES	1	STU	NO	3.31	0.0011
PSY	YES	1	STU	NO	6.30	<.0001
PSY	YES	1	STU	NO	9.02	<.0001
PSY	YES	1	STU	YES	-0.96	0.3405
PSY	YES	1	STU	YES	2.25	0.0256
PSY	YES	1	STU	YES	4.65	<.0001
PSY	YES	1	STU	YES	7.50	<.0001
PSY	YES	3	PSY	YES	4.13	<.0001
PSY	YES	3	PSY	YES	6.97	<.0001
PSY	YES	3	STU	NO	-1.49	0.1376

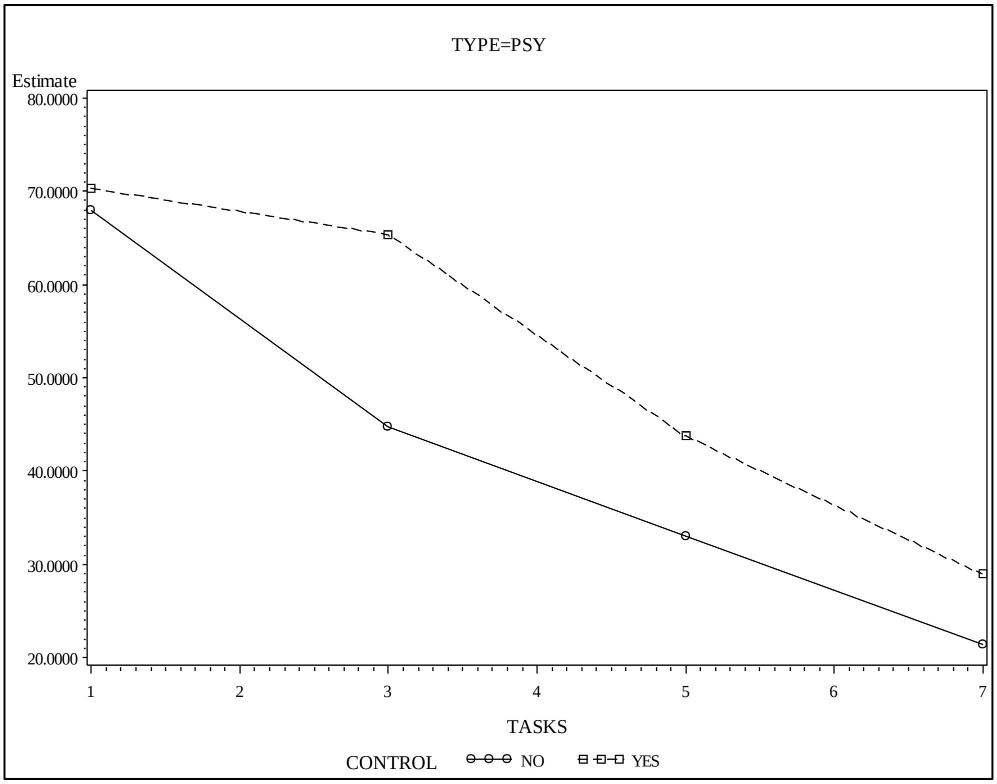
Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
PSY	YES	3	STU	NO	2.50	0.0133
PSY	YES	3	STU	NO	5.49	<.0001
PSY	YES	3	STU	NO	8.21	<.0001
PSY	YES	3	STU	YES	-1.77	0.0779
PSY	YES	3	STU	YES	1.43	0.1536
PSY	YES	3	STU	YES	3.84	0.0002
PSY	YES	3	STU	YES	6.68	<.0001
PSY	YES	5	PSY	YES	2.84	0.0057
PSY	YES	5	STU	NO	-5.04	<.0001
PSY	YES	5	STU	NO	-1.05	0.2959
PSY	YES	5	STU	NO	1.94	0.0536
PSY	YES	5	STU	NO	4.67	<.0001
PSY	YES	5	STU	YES	-5.35	<.0001
PSY	YES	5	STU	YES	-2.15	0.0329
PSY	YES	5	STU	YES	0.26	0.7975
PSY	YES	5	STU	YES	3.10	0.0022
PSY	YES	7	STU	NO	-7.47	<.0001
PSY	YES	7	STU	NO	-3.48	0.0006
PSY	YES	7	STU	NO	-0.50	0.6210
PSY	YES	7	STU	NO	2.23	0.0268
PSY	YES	7	STU	YES	-7.81	<.0001
PSY	YES	7	STU	YES	-4.61	<.0001
PSY	YES	7	STU	YES	-2.20	0.0287
PSY	YES	7	STU	YES	0.64	0.5203
STU	NO	1	STU	NO	10.92	<.0001
STU	NO	1	STU	NO	19.10	<.0001
STU	NO	1	STU	NO	26.56	<.0001
STU	NO	1	STU	YES	-0.41	0.6794
STU	NO	1	STU	YES	4.56	<.0001
STU	NO	1	STU	YES	8.29	<.0001
STU	NO	1	STU	YES	12.71	<.0001
STU	NO	3	STU	NO	8.18	<.0001

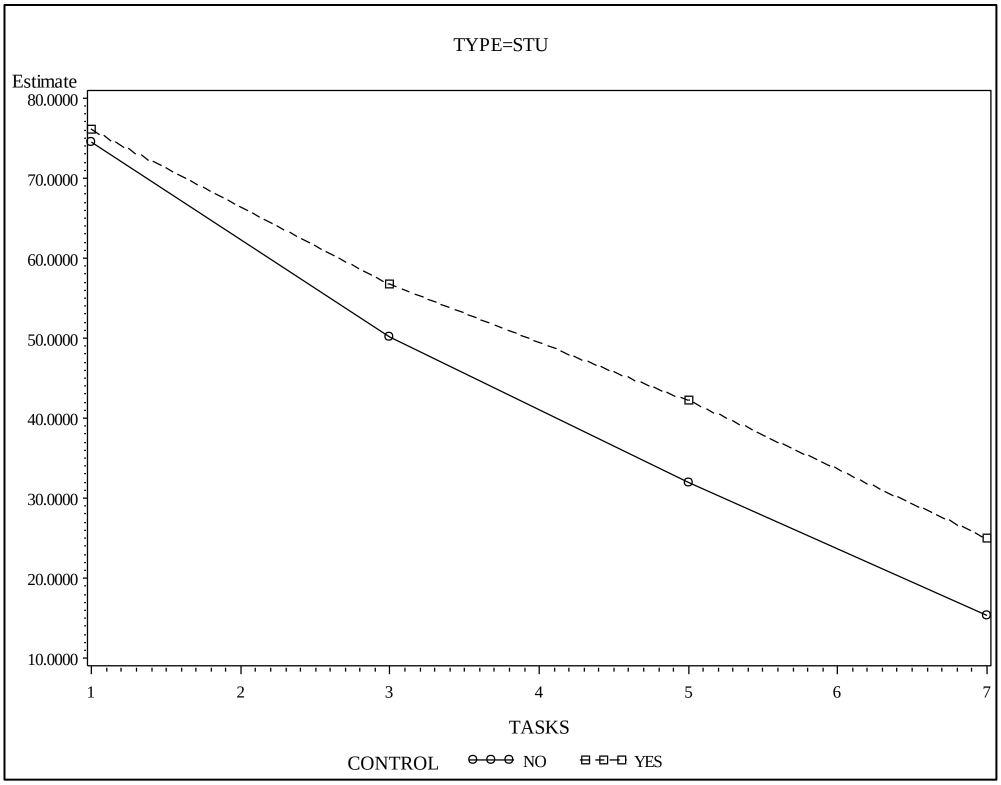
Differences of Least Squares Means						
TYPE	**CONTROL**	**TASKS**	**TYPE**	**CONTROL**	**t Value**	**Pr > |t|**
STU	NO	3	STU	NO	15.65	<.0001
STU	NO	3	STU	YES	-6.66	<.0001
STU	NO	3	STU	YES	-1.69	0.0926
STU	NO	3	STU	YES	2.04	0.0427
STU	NO	3	STU	YES	6.46	<.0001
STU	NO	5	STU	NO	7.46	<.0001
STU	NO	5	STU	YES	-11.35	<.0001
STU	NO	5	STU	YES	-6.38	<.0001
STU	NO	5	STU	YES	-2.64	0.0089
STU	NO	5	STU	YES	1.78	0.0775
STU	NO	7	STU	YES	-15.62	<.0001
STU	NO	7	STU	YES	-10.65	<.0001
STU	NO	7	STU	YES	-6.92	<.0001
STU	NO	7	STU	YES	-2.50	0.0134
STU	YES	1	STU	YES	9.11	<.0001
STU	YES	1	STU	YES	15.95	<.0001
STU	YES	1	STU	YES	24.04	<.0001
STU	YES	3	STU	YES	6.84	<.0001
STU	YES	3	STU	YES	14.93	<.0001
STU	YES	5	STU	YES	8.10	<.0001

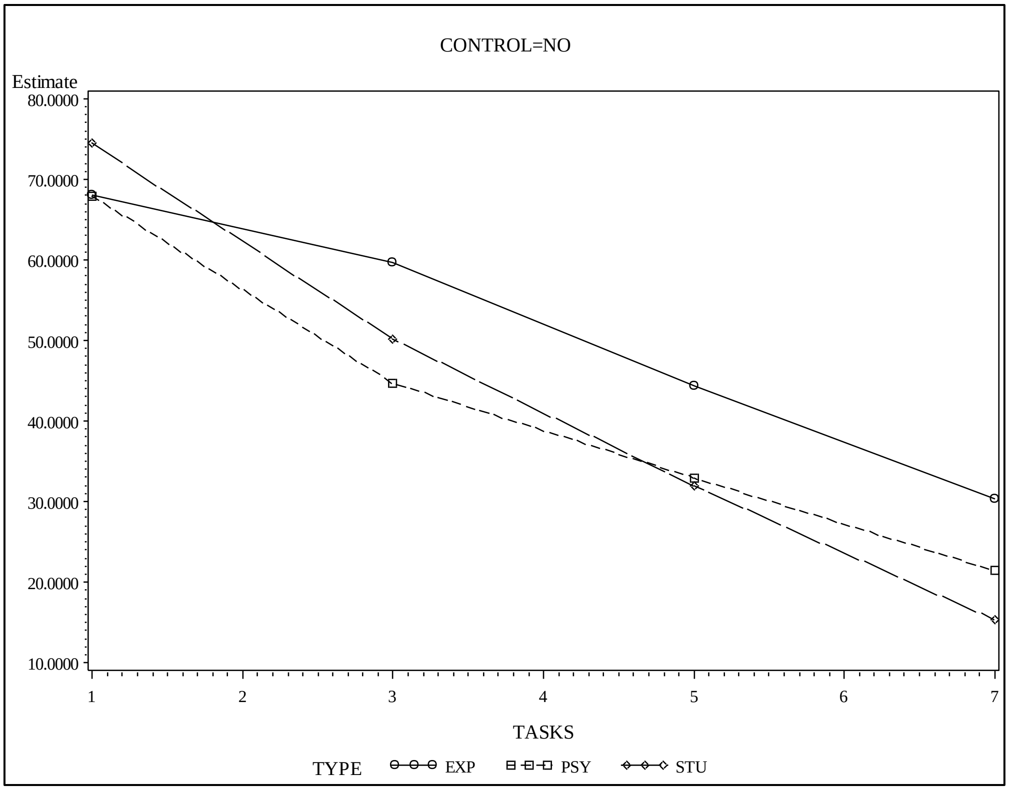
Tests of Effect Slices						
Effect	**TYPE**	**CONTROL**	**TASKS**	**Num DF**	**F Value**	**Pr > F**
TYPE*CONTROL*TASKS		NO	1	2	0.78	0.4593
TYPE*CONTROL*TASKS		NO	3	2	1.29	0.2767
TYPE*CONTROL*TASKS		NO	5	2	1.13	0.3257
TYPE*CONTROL*TASKS		NO	7	2	1.90	0.1529
TYPE*CONTROL*TASKS		YES	1	2	0.51	0.6011
TYPE*CONTROL*TASKS		YES	3	2	2.75	0.0661
TYPE*CONTROL*TASKS		YES	5	2	3.10	0.0472
TYPE*CONTROL*TASKS		YES	7	2	5.45	0.0049

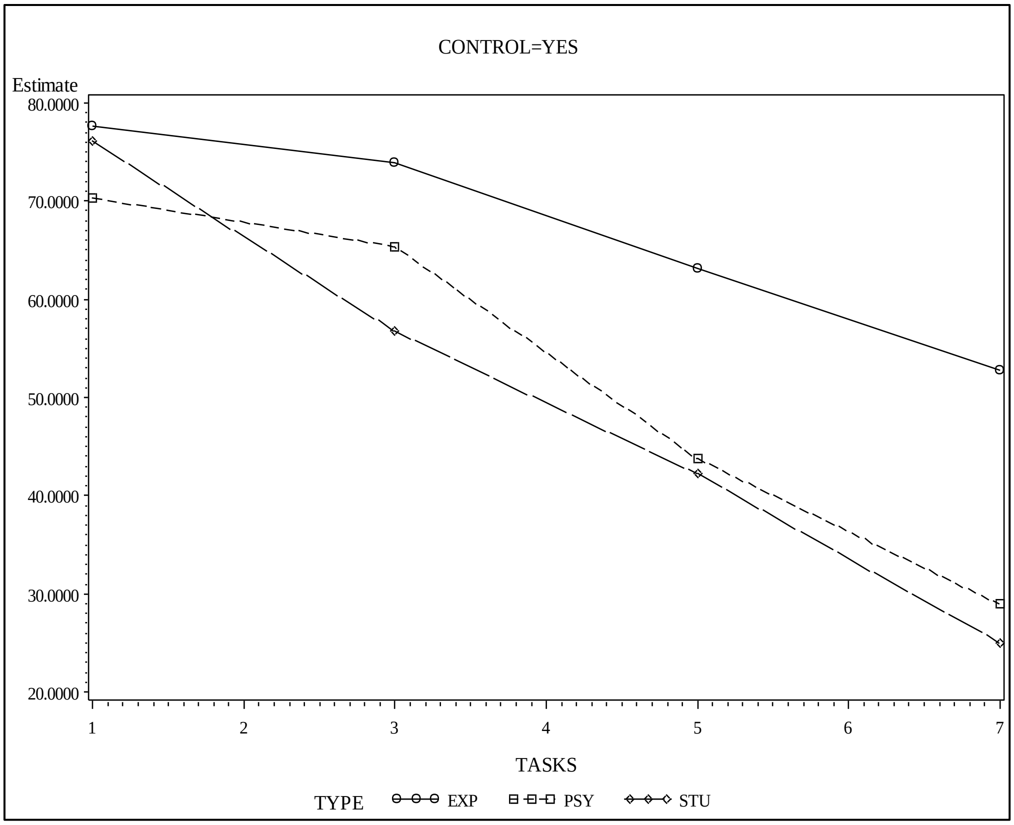
Tests of Effect Slices						
Effect	**TYPE**	**CONTROL**	**TASKS**	**Num DF**	**F Value**	**Pr > F**
TYPE*CONTROL*TASKS	EXP		1	1	0.74	0.3916
TYPE*CONTROL*TASKS	EXP		3	1	1.62	0.2053
TYPE*CONTROL*TASKS	EXP		5	1	2.83	0.0944
TYPE*CONTROL*TASKS	EXP		7	1	4.09	0.0448
TYPE*CONTROL*TASKS	PSY		1	1	0.10	0.7558
TYPE*CONTROL*TASKS	PSY		3	1	7.56	0.0066
TYPE*CONTROL*TASKS	PSY		5	1	2.07	0.1520
TYPE*CONTROL*TASKS	PSY		7	1	0.99	0.3201
TYPE*CONTROL*TASKS	STU		1	1	0.17	0.6794
TYPE*CONTROL*TASKS	STU		3	1	2.86	0.0926
TYPE*CONTROL*TASKS	STU		5	1	6.99	0.0089
TYPE*CONTROL*TASKS	STU		7	1	6.24	0.0134




TYPE=PSY
Estimate
80.0000
70.0000
60.0000
50.0000
40.0000
30.0000
20.0000
1
2
3
4
5
6
7
TASKS
CONTROL
NO
YES


TYPE=STU
Estimate
80.0000
70.0000
60.0000
50.0000
40.0000
30.0000
20.0000
10.0000
1
2
3
4
5
6
7
TASKS
CONTROL
NO
YES


CONTROL=NO
Estimate
80.0000
70.0000
60.0000
50.0000
40.0000
30.0000
20.0000
10.0000
1
2
3
4
5
6
7
TASKS
TYPE
EXP
PSY
STU


CONTROL=YES
Estimate
80.0000
70.0000
60.0000
50.0000
40.0000
30.0000
20.0000
1
2
3
4
5
6
7
TASKS
TYPE
EXP
PSY
STU

Model Information	
Data Set	WORK.ALL
Dependent Variable	CONFIDENCE
Covariance Structures	Variance Components, Unstructured
Group Effect	TYPE
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	**Levels**	**Values**
TYPE	3	EXP PSY STU
CONTROL	2	NO YES
TASKS	4	1 3 5 7

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
0	1	4629.42313897	
1	2	4365.89611566	0.00001803
2	1	4365.86472817	0.00000002
3	1	4365.86469279	0.00000000

Convergence criteria met.

Covariance Parameter Estimates		
Cov Parm	**Group**	**Estimate**
Student		217.47
UN(1,1)	TYPE EXP	351.49
UN(1,1)	TYPE PSY	191.19
UN(1,1)	TYPE STU	99.0268

Fit Statistics	
-2 Res Log Likelihood	4365.9
AIC (smaller is better)	4373.9
AICC (smaller is better)	4373.9
BIC (smaller is better)	4385.2

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
TASKS	3	139	137.88	<.0001
CONTROL	1	136	8.92	0.0033
CONTROL*TASKS	3	139	1.95	0.1240
TYPE	2	132	3.38	0.0369
TYPE*TASKS	6	98.1	4.24	0.0008
TYPE*CONTROL	2	132	0.52	0.5954
TYPE*CONTROL*TASKS	6	98.1	0.94	0.4683

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**EXP TASKS 1 VS 3**	6.0216	5.5337	71.42	1.09	0.2802

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**EXP TASKS 3 VS 5**	13.0481	5.5337	71.42	2.36	0.0211

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**EXP TASKS 5 VS 7**	12.2481	5.5337	71.42	2.21	0.0301

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**PSY TASKS 1 VS 3**	14.0862	3.6333	81.76	3.88	0.0002

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**PSY TASKS 3 VS 5**	16.6675	3.6333	81.76	4.59	<.0001

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**PSY TASKS 5 VS 7**	13.1848	3.6333	81.76	3.63	0.0005

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**STU TASKS 1 VS 3**	21.8110	1.5373	246.7	14.19	<.0001

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**STU TASKS 3 VS 5**	16.3573	1.5373	246.7	10.64	<.0001

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**STU TASKS 5 VS 7**	16.8911	1.5373	246.7	10.99	<.0001

Least Squares Means							
Effect	**TYPE**	**TASKS**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	EXP	1	72.8624	5.5829	166	13.05	<.0001
TYPE*TASKS	EXP	3	66.8408	5.5829	166	11.97	<.0001
TYPE*TASKS	EXP	5	53.7927	5.5829	166	9.64	<.0001
TYPE*TASKS	EXP	7	41.5446	5.5829	166	7.44	<.0001
TYPE*TASKS	PSY	1	69.1300	3.7561	181	18.40	<.0001
TYPE*TASKS	PSY	3	55.0438	3.7561	181	14.65	<.0001
TYPE*TASKS	PSY	5	38.3763	3.7561	181	10.22	<.0001
TYPE*TASKS	PSY	7	25.1915	3.7561	181	6.71	<.0001
TYPE*TASKS	STU	1	75.2593	1.9433	183	38.73	<.0001
TYPE*TASKS	STU	3	53.4483	1.9433	183	27.50	<.0001

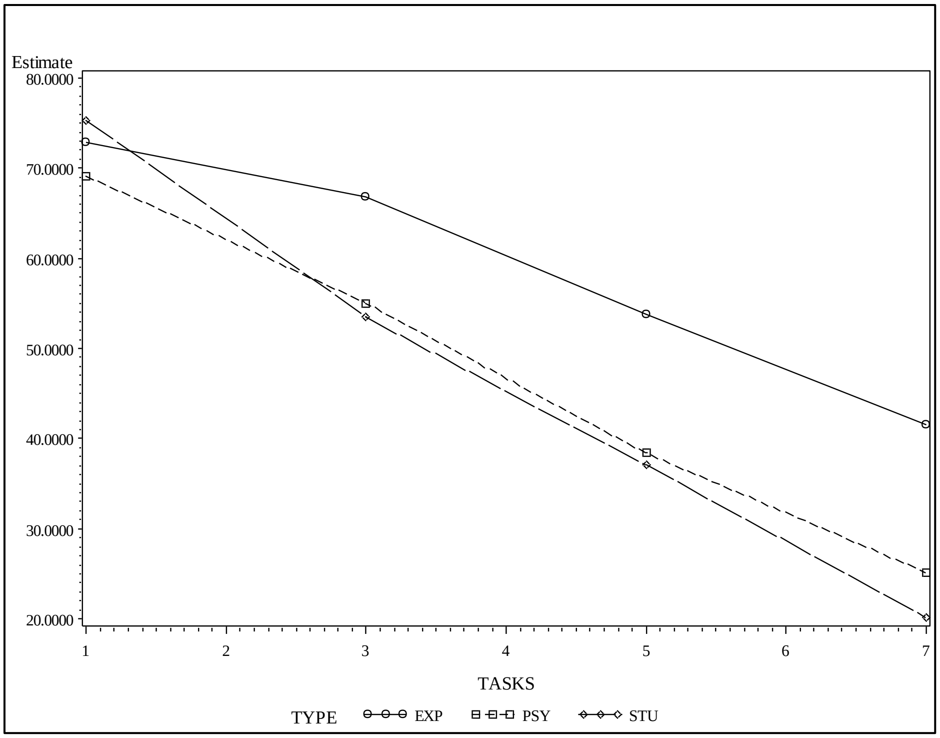
Least Squares Means							
Effect	**TYPE**	**TASKS**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	STU	5	37.0910	1.9433	183	19.09	<.0001
TYPE*TASKS	STU	7	20.1999	1.9433	183	10.39	<.0001

Differences of Least Squares Means								
Effect	**TYPE**	**TASKS**	**TYPE**	**TASKS**	**Estimate**	**S Error**	**t Value**	**Pr > |t|**
TYPE*TASKS	EXP	1	EXP	3	6.0216	5.5337	1.09	0.2802
TYPE*TASKS	EXP	1	EXP	5	19.0697	5.5337	3.45	0.0010
TYPE*TASKS	EXP	1	EXP	7	31.3178	5.5337	5.66	<.0001
TYPE*TASKS	EXP	1	PSY	1	3.7324	6.7288	0.55	0.5797
TYPE*TASKS	EXP	1	PSY	3	17.8186	6.7288	2.65	0.0087
TYPE*TASKS	EXP	1	PSY	5	34.4861	6.7288	5.13	<.0001
TYPE*TASKS	EXP	1	PSY	7	47.6709	6.7288	7.08	<.0001
TYPE*TASKS	EXP	1	STU	1	-2.3969	5.9114	-0.41	0.6856
TYPE*TASKS	EXP	1	STU	3	19.4141	5.9114	3.28	0.0012
TYPE*TASKS	EXP	1	STU	5	35.7714	5.9114	6.05	<.0001
TYPE*TASKS	EXP	1	STU	7	52.6625	5.9114	8.91	<.0001
TYPE*TASKS	EXP	3	EXP	5	13.0481	5.5337	2.36	0.0211
TYPE*TASKS	EXP	3	EXP	7	25.2962	5.5337	4.57	<.0001
TYPE*TASKS	EXP	3	PSY	1	-2.2892	6.7288	-0.34	0.7340
TYPE*TASKS	EXP	3	PSY	3	11.7970	6.7288	1.75	0.0810
TYPE*TASKS	EXP	3	PSY	5	28.4645	6.7288	4.23	<.0001
TYPE*TASKS	EXP	3	PSY	7	41.6493	6.7288	6.19	<.0001
TYPE*TASKS	EXP	3	STU	1	-8.4185	5.9114	-1.42	0.1561
TYPE*TASKS	EXP	3	STU	3	13.3925	5.9114	2.27	0.0247
TYPE*TASKS	EXP	3	STU	5	29.7498	5.9114	5.03	<.0001
TYPE*TASKS	EXP	3	STU	7	46.6409	5.9114	7.89	<.0001
TYPE*TASKS	EXP	5	EXP	7	12.2481	5.5337	2.21	0.0301
TYPE*TASKS	EXP	5	PSY	1	-15.3373	6.7288	-2.28	0.0236
TYPE*TASKS	EXP	5	PSY	3	-1.2511	6.7288	-0.19	0.8527
TYPE*TASKS	EXP	5	PSY	5	15.4164	6.7288	2.29	0.0229

Differences of Least Squares Means								
Effect	**TYPE**	**TASKS**	**TYPE**	**TASKS**	**Estimate**	**S Error**	**t Value**	**Pr > |t|**
TYPE*TASKS	EXP	5	PSY	7	28.6012	6.7288	4.25	<.0001
TYPE*TASKS	EXP	5	STU	1	-21.4666	5.9114	-3.63	0.0004
TYPE*TASKS	EXP	5	STU	3	0.3444	5.9114	0.06	0.9536
TYPE*TASKS	EXP	5	STU	5	16.7017	5.9114	2.83	0.0053
TYPE*TASKS	EXP	5	STU	7	33.5928	5.9114	5.68	<.0001
TYPE*TASKS	EXP	7	PSY	1	-27.5854	6.7288	-4.10	<.0001
TYPE*TASKS	EXP	7	PSY	3	-13.4992	6.7288	-2.01	0.0461
TYPE*TASKS	EXP	7	PSY	5	3.1683	6.7288	0.47	0.6382
TYPE*TASKS	EXP	7	PSY	7	16.3531	6.7288	2.43	0.0159
TYPE*TASKS	EXP	7	STU	1	-33.7147	5.9114	-5.70	<.0001
TYPE*TASKS	EXP	7	STU	3	-11.9037	5.9114	-2.01	0.0455
TYPE*TASKS	EXP	7	STU	5	4.4536	5.9114	0.75	0.4522
TYPE*TASKS	EXP	7	STU	7	21.3447	5.9114	3.61	0.0004
TYPE*TASKS	PSY	1	PSY	3	14.0862	3.6333	3.88	0.0002
TYPE*TASKS	PSY	1	PSY	5	30.7537	3.6333	8.46	<.0001
TYPE*TASKS	PSY	1	PSY	7	43.9385	3.6333	12.09	<.0001
TYPE*TASKS	PSY	1	STU	1	-6.1293	4.2290	-1.45	0.1488
TYPE*TASKS	PSY	1	STU	3	15.6817	4.2290	3.71	0.0003
TYPE*TASKS	PSY	1	STU	5	32.0390	4.2290	7.58	<.0001
TYPE*TASKS	PSY	1	STU	7	48.9301	4.2290	11.57	<.0001
TYPE*TASKS	PSY	3	PSY	5	16.6675	3.6333	4.59	<.0001
TYPE*TASKS	PSY	3	PSY	7	29.8523	3.6333	8.22	<.0001
TYPE*TASKS	PSY	3	STU	1	-20.2155	4.2290	-4.78	<.0001
TYPE*TASKS	PSY	3	STU	3	1.5955	4.2290	0.38	0.7064
TYPE*TASKS	PSY	3	STU	5	17.9528	4.2290	4.25	<.0001
TYPE*TASKS	PSY	3	STU	7	34.8439	4.2290	8.24	<.0001
TYPE*TASKS	PSY	5	PSY	7	13.1848	3.6333	3.63	0.0005
TYPE*TASKS	PSY	5	STU	1	-36.8830	4.2290	-8.72	<.0001
TYPE*TASKS	PSY	5	STU	3	-15.0720	4.2290	-3.56	0.0005
TYPE*TASKS	PSY	5	STU	5	1.2853	4.2290	0.30	0.7615

Differences of Least Squares Means								
Effect	**TYPE**	**TASKS**	**TYPE**	**TASKS**	**Estimate**	**S Error**	**t Value**	**Pr > |t|**
TYPE*TASKS	PSY	5	STU	7	18.1764	4.2290	4.30	<.0001
TYPE*TASKS	PSY	7	STU	1	-50.0678	4.2290	-11.84	<.0001
TYPE*TASKS	PSY	7	STU	3	-28.2567	4.2290	-6.68	<.0001
TYPE*TASKS	PSY	7	STU	5	-11.8995	4.2290	-2.81	0.0054
TYPE*TASKS	PSY	7	STU	7	4.9917	4.2290	1.18	0.2393
TYPE*TASKS	STU	1	STU	3	21.8110	1.5373	14.19	<.0001
TYPE*TASKS	STU	1	STU	5	38.1683	1.5373	24.83	<.0001
TYPE*TASKS	STU	1	STU	7	55.0594	1.5373	35.82	<.0001
TYPE*TASKS	STU	3	STU	5	16.3573	1.5373	10.64	<.0001
TYPE*TASKS	STU	3	STU	7	33.2484	1.5373	21.63	<.0001
TYPE*TASKS	STU	5	STU	7	16.8911	1.5373	10.99	<.0001

Tests of Effect Slices						
Effect	**TYPE**	**TASKS**	**Num DF**	**Den DF**	**F Value**	**Pr > F**
TYPE*TASKS		1	2	201	1.07	0.3456
TYPE*TASKS		3	2	201	2.57	0.0793
TYPE*TASKS		5	2	201	4.00	0.0198
TYPE*TASKS		7	2	201	6.73	0.0015
TYPE*TASKS	EXP		3	71.4	12.74	<.0001
TYPE*TASKS	PSY		3	81.8	55.77	<.0001
TYPE*TASKS	STU		3	247	467.06	<.0001



Model Information	
Data Set	WORK.ALL
Dependent Variable	CONFIDENCE
Covariance Structures	Variance Components, Unstructured
Group Effect	TYPE
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	**Levels**	**Values**
TYPE	3	EXP PSY STU
CONTROL	2	NO YES
TASKS	4	1 3 5 7

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
0	1	4629.42313897	
1	2	4365.89611566	0.00001803
2	1	4365.86472817	0.00000002
3	1	4365.86469279	0.00000000

Convergence criteria met.

Covariance Parameter Estimates		
Cov Parm	**Group**	**Estimate**
Student		217.47
UN(1,1)	TYPE EXP	351.49
UN(1,1)	TYPE PSY	191.19
UN(1,1)	TYPE STU	99.0268

Fit Statistics	
-2 Res Log Likelihood	4365.9
AIC (smaller is better)	4373.9
AICC (smaller is better)	4373.9
BIC (smaller is better)	4385.2

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
TASKS	3	139	137.88	<.0001
CONTROL	1	136	8.92	0.0033
CONTROL*TASKS	3	139	1.95	0.1240
TYPE	2	132	3.38	0.0369
TYPE*TASKS	6	98.1	4.24	0.0008
TYPE*CONTROL	2	132	0.52	0.5954
TYPE*CONTROL*TASKS	6	98.1	0.94	0.4683

Least Squares Means							
Effect	**CONTROL**	**TASKS**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
CONTROL*TASKS	NO	1	70.1605	3.2671	225	21.47	<.0001
CONTROL*TASKS	NO	3	51.5391	3.2671	225	15.77	<.0001
CONTROL*TASKS	NO	5	36.4419	3.2671	225	11.15	<.0001
CONTROL*TASKS	NO	7	22.3508	3.2671	225	6.84	<.0001
CONTROL*TASKS	YES	1	74.6740	3.3358	227	22.39	<.0001
CONTROL*TASKS	YES	3	65.3495	3.3358	227	19.59	<.0001
CONTROL*TASKS	YES	5	49.7315	3.3358	227	14.91	<.0001
CONTROL*TASKS	YES	7	35.6066	3.3358	227	10.67	<.0001

Tests of Effect Slices						
Effect	**CONTROL**	**TASKS**	**Num DF**	**Den DF**	**F Value**	**Pr > F**
CONTROL*TASKS		1	1	226	0.93	0.3348
CONTROL*TASKS		3	1	226	8.75	0.0034
CONTROL*TASKS		5	1	226	8.10	0.0048
CONTROL*TASKS		7	1	226	8.06	0.0049
CONTROL*TASKS	NO		3	141	85.05	<.0001
CONTROL*TASKS	YES		3	138	55.84	<.0001

The below mentioned SAS program is used to combine two population types and compare it with a single population.

```
ODS RTF FILE=’U: \test2.rtf’;
DATA NOCONTROLST;
INPUT Student $ T1 T3 T5 T7;
TYPE='STU';
Student = student+100;
DATALINES;
1       66.3    54.3    7.6     0
2       94.6    71.7    19.6    13
3       75      63.8    47.4    13.2
.       .       .       .       .
.       .       .       .       .
.       .       .       .       .
;
DATA T1;
SET NOCONTROLST;
CONFIDENCE=T1;
```

```
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET NOCONTROLST;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET NOCONTROLST;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET NOCONTROLST;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA NCST;
SET T1 T3 T5 T7;
CONTROL='NO ';
*PROC PRINT;
RUN;
QUIT;

DATA CONTROLST;
INPUT Student $ T1 T3 T5 T7;
TYPE='STU';
Student = student+200;
DATALINES;
1         22.8     13       12       7.6
2         96.7     76.1     70.7     50
.         .        .        .        .
.         .        .        .        .
.         .        .        .        .
;

DATA T1;
SET CONTROLST;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET CONTROLST;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET CONTROLST;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET CONTROLST;
```

```
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA WCST;
SET T1 T3 T5 T7;
CONTROL='YES';

DATA NOCONTROLPSY;
INPUT Student $ T1 T3 T5 T7;
TYPE='PSY';
Student = student+300;
DATALINES;
1         63.4      39.8      18.3      9.7
2         78.5      61.3      24.7      6.5
3         45.2      32.3      16.1      8.6
.         .         .         .         .
.         .         .         .         .
.         .         .         .         .
;
DATA T1;
SET NOCONTROLPSY;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET NOCONTROLPSY;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET NOCONTROLPSY;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET NOCONTROLPSY;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA NCPSY;
SET T1 T3 T5 T7;
CONTROL='NO ';
*PROC PRINT;
RUN;
QUIT;

DATA CONTROLPSY;
INPUT Student $ T1 T3 T5 T7;
TYPE='PSY';
Student =student+400;
DATALINES;
1         87        79.6      77.4      76.3
2         52.7      40        20.4      9.6
```

```
.          .          .          .          .
.          .          .          .          .
.          .          .          .          .
;
DATA T1;
SET CONTROLPSY;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET CONTROLPSY;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET CONTROLPSY;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET CONTROLPSY;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA WCPSY;
SET T1 T3 T5 T7;
CONTROL='YES';

DATA NOCONTROLEXP;
INPUT Student $ T1 T3 T5 T7;
TYPE='EXP';
Student = student+500;
DATALINES;
1          92.4       89.1       73.9       56.5
2          43.5       32.6       61         25
.          .          .          .          .
.          .          .          .          .
.          .          .          .          .
;
DATA T1;
SET NOCONTROLEXP;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET NOCONTROLEXP;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET NOCONTROLEXP;
CONFIDENCE=T5;
TASKS=5;
```

```
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET NOCONTROLEXP;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA NCEXP;
SET T1 T3 T5 T7;
CONTROL='NO ';
*PROC PRINT;
RUN;
QUIT;

DATA CONTROLEXP;
INPUT Student $ T1 T3 T5 T7;
TYPE='EXP';
Student =student+600;
DATALINES;
1       22.8    3.2     28.3    0
2       98      97      95      90
.       .       .       .       .
.       .       .       .       .
.       .       .       .       .
;
DATA T1;
SET CONTROLEXP;
CONFIDENCE=T1;
TASKS=1;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T3;
SET CONTROLEXP;
CONFIDENCE=T3;
TASKS=3;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T5;
SET CONTROLEXP;
CONFIDENCE=T5;
TASKS=5;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA T7;
SET CONTROLEXP;
CONFIDENCE=T7;
TASKS=7;
KEEP TYPE STUDENT CONFIDENCE TASKS;
DATA WCEXP;
SET T1 T3 T5 T7;
CONTROL='YES';
DATA ALL;
SET NCST WCST ncpsy wcpsy ncexp wcexp;

PROC MIXED DATA=ALL;
CLASS TYPE CONTROL TASKS STUDENT;
MODEL CONFIDENCE=TASKS|CONTROL|TYPE /OUTP=RESIDUAL DDFM=SAT;
```

```
RANDOM STUDENT;
REPEATED / TYPE=UN GROUP=TYPE;
LSMEANS TASKS*TYPE /slice=type slice=tasks;
LSMEANS TASKS*TYPE*CONTROL;

LSMESTIMATE TASKS*TYPE 'EXP VS OTHERS TASKS 1' 1 0 0 0 -.5 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE 'EXP VS OTHERS TASKS 3' 0 1 0 0 0 -.5 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE 'EXP VS OTHERS TASKS 5' 0 0 1 0 0 0 -.5 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE 'EXP VS OTHERS TASKS 7' 0 0 0 1 0 0 0 -.5 0 0 0 -.5;

LSMESTIMATE TASKS*TYPE*CONTROL 'EXP VS OTHERS TASKS 1 NO CONTROL' 1 0 0 0 0
0 0 0 -.5 0 0 0 0 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE*CONTROL 'EXP VS OTHERS TASKS 3 NO CONTROL' 0 1 0 0 0
0 0 0 0 -.5 0 0 0 0 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE*CONTROL 'EXP VS OTHERS TASKS 5 NO CONTROL' 0 0 1 0 0
0 0 0 0 0 -.5 0 0 0 0 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE*CONTROL 'EXP VS OTHERS TASKS 7 NO CONTROL' 0 0 0 1 0
0 0 0 0 0 0 -.5 0 0 0 0 0 0 0 -.5;

LSMESTIMATE TASKS*TYPE*CONTROL 'EXP VS OTHERS TASKS 1 CONTROL' 0 0 0 0 1 0 0
0 0 0 0 0 -.5 0 0 0 0 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE*CONTROL 'EXP VS OTHERS TASKS 3 CONTROL' 0 0 0 0 0 1 0
0 0 0 0 0 0 -.5 0 0 0 0 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE*CONTROL 'EXP VS OTHERS TASKS 5 CONTROL' 0 0 0 0 0 0 1
0 0 0 0 0 0 0 -.5 0 0 0 0 0 0 0 -.5;
LSMESTIMATE TASKS*TYPE*CONTROL 'EXP VS OTHERS TASKS 7 CONTROL' 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 -.5 0 0 0 0 0 0 0 -.5;


LSMESTIMATE TASKS*TYPE 'PSY VS OTHERS TASKS 1' .5 0 0 0 -1 0 0 0 .5;
LSMESTIMATE TASKS*TYPE 'PSY VS OTHERS TASKS 3' 0 .5 0 0 0 -1 0 0 0 .5;
LSMESTIMATE TASKS*TYPE 'PSY VS OTHERS TASKS 5' 0 0 .5 0 0 0 -1 0 0 0 .5;
LSMESTIMATE TASKS*TYPE 'PSY VS OTHERS TASKS 7' 0 0 0 .5 0 0 0 -1 0 0 0 .5;

LSMESTIMATE TASKS*TYPE*CONTROL 'PSY VS OTHERS TASKS 1 NO CONTROL' .5 0 0 0
0 0 0 0 -1 0 0 0 0 0 0 0 .5;
LSMESTIMATE TASKS*TYPE*CONTROL 'PSY VS OTHERS TASKS 3 NO CONTROL' 0 .5 0 0
0 0 0 0 0 -1 0 0 0 0 0 0 0 .5;
LSMESTIMATE TASKS*TYPE*CONTROL 'PSY VS OTHERS TASKS 5 NO CONTROL' 0 0 .5 0
0 0 0 0 0 0 -1 0 0 0 0 0 0 0 .5;
LSMESTIMATE TASKS*TYPE*CONTROL 'PSY VS OTHERS TASKS 7 NO CONTROL' 0 0 0 .5
0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 .5;

LSMESTIMATE TASKS*TYPE*CONTROL 'PSY VS OTHERS TASKS 1 CONTROL' 0 0 0 0 .5 0
0 0 0 0 0 0 -1 0 0 0 0 0 0 0 .5;
LSMESTIMATE TASKS*TYPE*CONTROL 'PSY VS OTHERS TASKS 3 CONTROL' 0 0 0 0 0 .5
0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 .5;
LSMESTIMATE TASKS*TYPE*CONTROL 'PSY VS OTHERS TASKS 5 CONTROL' 0 0 0 0 0 0
.5 0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 .5;
LSMESTIMATE TASKS*TYPE*CONTROL 'PSY VS OTHERS TASKS 7 CONTROL' 0 0 0 0 0 0 0
.5 0 0 0 0 0 0 0 -1 0 0 0 0 0 0 0 .5;
ODS RTF CLOSE;
RUN;
```

```
QUIT;
```

OUTPUT:

Model Information	
Data Set	WORK.ALL
Dependent Variable	CONFIDENCE
Covariance Structures	Variance Components, Unstructured
Group Effect	TYPE
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Satterthwaite

Class Level Information		
Class	**Levels**	**Values**
TYPE	3	EXP PSY STU
CONTROL	2	NO YES
TASKS	4	1 3 5 7

Iteration History			
Iteration	**Evaluations**	**-2 Res Log Like**	**Criterion**
0	1	4629.42313897	
1	2	4365.89611566	0.00001803
2	1	4365.86472817	0.00000002
3	1	4365.86469279	0.00000000

Convergence criteria met.

Covariance Parameter Estimates		
Cov Parm	**Group**	**Estimate**
Student		217.47
UN(1,1)	TYPE EXP	351.49
UN(1,1)	TYPE PSY	191.19
UN(1,1)	TYPE STU	99.0268

Fit Statistics	
-2 Res Log Likelihood	4365.9
AIC (smaller is better)	4373.9
AICC (smaller is better)	4373.9
BIC (smaller is better)	4385.2

Type 3 Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
TASKS	3	139	137.88	<.0001
CONTROL	1	136	8.92	0.0033
CONTROL*TASKS	3	139	1.95	0.1240
TYPE	2	132	3.38	0.0369
TYPE*TASKS	6	98.1	4.24	0.0008
TYPE*CONTROL	2	132	0.52	0.5954
TYPE*CONTROL*TASKS	6	98.1	0.94	0.4683

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**EXP VS OTHERS TASKS 1**	0.6678	5.9699	188	0.11	0.9111

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**EXP VS OTHERS TASKS 3**	12.5948	5.9699	188	2.11	0.0362

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**EXP VS OTHERS TASKS 5**	16.0591	5.9699	188	2.69	0.0078

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASKS	**EXP VS OTHERS TASKS 7**	18.8489	5.9699	188	3.16	0.0019

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
EXP VS OTHERS TASKS 1 NO CONTROL	-3.1410	8.3487	188.3	-0.38	0.7072

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
EXP VS OTHERS TASKS 3 NO CONTROL	12.3036	8.3487	188.3	1.47	0.1422

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
EXP VS OTHERS TASKS 5 NO CONTROL	11.9369	8.3487	188.3	1.43	0.1544

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
EXP VS OTHERS TASKS 7 NO CONTROL	11.8611	8.3487	188.3	1.42	0.1571

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
EXP VS OTHERS TASKS 1 CONTROL	4.4766	8.5357	187.5	0.52	0.6006

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
EXP VS OTHERS TASKS 3 CONTROL	12.8860	8.5357	187.5	1.51	0.1328

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
EXP VS OTHERS TASKS 5 CONTROL	20.1812	8.5357	187.5	2.36	0.0191

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
EXP VS OTHERS TASKS 7 CONTROL	25.8367	8.5357	187.5	3.03	0.0028

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASK	**PSY VS OTHERS TASKS 1**	4.9309	4.7796	234.9	1.03	0.3033

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASK	**PSY VS OTHERS TASKS 3**	5.1008	4.7796	234.9	1.07	0.2870

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASK	**PSY VS OTHERS TASKS 5**	7.0656	4.7796	234.9	1.48	0.1407

Least Squares Means Estimate						
Effect	**Label**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASK	**PSY VS OTHERS TASKS 7**	5.6807	4.7796	234.9	1.19	0.2358

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
PSY VS OTHERS TASKS 1 NO CONTROL	3.3007	6.6671	233.8	0.50	0.6210

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
PSY VS OTHERS TASKS 3 NO CONTROL	10.2387	6.6671	233.8	1.54	0.1260

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
PSY VS OTHERS TASKS 5 NO CONTROL	5.2028	6.6671	233.8	0.78	0.4360

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
PSY VS OTHERS TASKS 7 NO CONTROL	1.3562	6.6671	233.8	0.20	0.8390

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
PSY VS OTHERS TASKS 1 CONTROL	6.5610	6.8504	236	0.96	0.3392

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
PSY VS OTHERS TASKS 3 CONTROL	-0.03715	6.8504	236	-0.01	0.9957

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
PSY VS OTHERS TASKS 5 CONTROL	8.9283	6.8504	236	1.30	0.1937

Least Squares Means Estimate					
Label	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
PSY VS OTHERS TASKS 7 CONTROL	10.0053	6.8504	236	1.46	0.1455

Least Squares Means								
Effect	**TYPE**	**CTRL**	**TASK**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASK	EXP		1	72.8624	5.5829	166	13.05	<.0001
TYPE*TASK	EXP		3	66.8408	5.5829	166	11.97	<.0001
TYPE*TASK	EXP		5	53.7927	5.5829	166	9.64	<.0001

Least Squares Means								
Effect	**TYPE**	**CTRL**	**TASK**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*TASK	EXP		7	41.5446	5.5829	166	7.44	<.0001
TYPE*TASK	PSY		1	69.1300	3.7561	181	18.40	<.0001
TYPE*TASK	PSY		3	55.0438	3.7561	181	14.65	<.0001
TYPE*TASK	PSY		5	38.3763	3.7561	181	10.22	<.0001
TYPE*TASK	PSY		7	25.1915	3.7561	181	6.71	<.0001
TYPE*TASK	STU		1	75.2593	1.9433	183	38.73	<.0001
TYPE*TASK	STU		3	53.4483	1.9433	183	27.50	<.0001
TYPE*TASK	STU		5	37.0910	1.9433	183	19.09	<.0001
TYPE*TASK	STU		7	20.1999	1.9433	183	10.39	<.0001
TYPE*CONTROL*TASKS	EXP	NO	1	68.0665	7.8046	167	8.72	<.0001
TYPE*CONTROL*TASKS	EXP	NO	3	59.7415	7.8046	167	7.65	<.0001
TYPE*CONTROL*TASKS	EXP	NO	5	44.3998	7.8046	167	5.69	<.0001
TYPE*CONTROL*TASKS	EXP	NO	7	30.2581	7.8046	167	3.88	0.0002
TYPE*CONTROL*TASKS	EXP	YES	1	77.6584	7.9852	165	9.73	<.0001
TYPE*CONTROL*TASKS	EXP	YES	3	73.9402	7.9852	165	9.26	<.0001
TYPE*CONTROL*TASKS	EXP	YES	5	63.1856	7.9852	165	7.91	<.0001
TYPE*CONTROL*TASKS	EXP	YES	7	52.8311	7.9852	165	6.62	<.0001
TYPE*CONTROL*TASKS	PSY	NO	1	67.9600	5.2195	181	13.02	<.0001
TYPE*CONTROL*TASKS	PSY	NO	3	44.7133	5.2195	181	8.57	<.0001
TYPE*CONTROL*TASKS	PSY	NO	5	32.9733	5.2195	181	6.32	<.0001
TYPE*CONTROL*TASKS	PSY	NO	7	21.4467	5.2195	181	4.11	<.0001

Least Squares Means								
Effect	**TYPE**	**CTRL**	**TASK**	**Estimate**	**S Error**	**DF**	**t Value**	**Pr > |t|**
TYPE*CONTROL*TASKS	PSY	YES	1	70.3000	5.4027	181	13.01	<.0001
TYPE*CONTROL*TASKS	PSY	YES	3	65.3743	5.4027	181	12.10	<.0001
TYPE*CONTROL*TASKS	PSY	YES	5	43.7793	5.4027	181	8.10	<.0001
TYPE*CONTROL*TASKS	PSY	YES	7	28.9364	5.4027	181	5.36	<.0001
TYPE*CONTROL*TASKS	STU	NO	1	74.4550	2.8129	183	26.47	<.0001
TYPE*CONTROL*TASKS	STU	NO	3	50.1625	2.8129	183	17.83	<.0001
TYPE*CONTROL*TASKS	STU	NO	5	31.9525	2.8129	183	11.36	<.0001
TYPE*CONTROL*TASKS	STU	NO	7	15.3475	2.8129	183	5.46	<.0001
TYPE*CONTROL*TASKS	STU	YES	1	76.0636	2.6820	183	28.36	<.0001
TYPE*CONTROL*TASKS	STU	YES	3	56.7341	2.6820	183	21.15	<.0001
TYPE*CONTROL*TASKS	STU	YES	5	42.2295	2.6820	183	15.75	<.0001
TYPE*CONTROL*TASKS	STU	YES	7	25.0523	2.6820	183	9.34	<.0001

Tests of Effect Slices						
Effect	**TYPE**	**TASKS**	**Num DF**	**Den DF**	**F Value**	**Pr > F**
TYPE*TASKS	EXP		3	71.4	12.74	<.0001
TYPE*TASKS	PSY		3	81.8	55.77	<.0001
TYPE*TASKS	STU		3	247	467.06	<.0001
TYPE*TASKS		1	2	201	1.07	0.3456
TYPE*TASKS		3	2	201	2.57	0.0793
TYPE*TASKS		5	2	201	4.00	0.0198
TYPE*TASKS		7	2	201	6.73	0.0015

REFERENCES

Arthur, J. D., & Groner, M. K. (2004). An operational model for structuring the requirements generation process. *Requirements Engineering*, *10*(1), 45–62. doi:10.1007/s00766-004-0196-2

Barnett, & Lewis. (1995). *Outliers in Statistical Data* (3rd Editio.). John Wiley & Sons. doi:10.1002/bimj.4710370219

Bonnabry, P., Despont-Gros, C., & Grauser, D. (2008). A Risk Analysis Method to Evaluate the Impact of a Computerized Provider Order Entry System on Patient Safety. *Journal of the American Medical Informatics Association*, *14*(4), 453–460.

Chapman, R. J. (1998). The role of system dynamics in understanding the impact of changes to key project personnel on design production within construction projects. *International Journal of Project Management*, *16*(4), 235–247. doi:10.1016/S0263-7863(97)00043-4

Clarkson, J., & Eckert, C. (2004). *Design Process Improvement: A Review of Current Practice*. Springer-Verlag London Limited.

Collins, S. T., Yassine, A. A., & Borgatti, S. P. (2009). Evaluating product development systems using network analysis. *Systems Engineering*, *12*(1), 55–68.

Corporation, C. (1995). *Potential Failure Mode and Effects Analysis Reference Manual*.

Cotnareanu, T. (1999). Old Tools-New Uses: Equipment FMEA, A Tool for Preventive Maintainance. *Quality Progress*, *32*(12), 48–52.

Davis, Stanley, William Riley, Ayse P. Gurses, Kristi Miller, and H. H. "Failure modes and effects analysis based on in situ simulations: a methodology to improve understanding of risks and failures. . A. in P. S. N. D. and A. A. 3 (2008): 145-160. (2008). Davis, Stanley, William Riley, Ayse P. Gurses, Kristi Miller, and Helen Hansen. “Failure modes and effects analysis based on in situ simulations: a methodology to improve understanding of risks and failures.” Advances in Patient Safety: New Directions and. *Advances in Patient Safety: New Directions and Alternative Approaches 3*, 145–160.

Department of Defense. (1980). Procedures for Performing A Failure Mode, Effects and Criticality Analysis. *MIL-STD-1629A*.

Dinsmore, D. L., & Parkinson, M. M. (2012). What are confidence judgments made of? Students' explanations for their confidence ratings and what that means for calibration. *Learning and Instruction*, 1–11. doi:10.1016/j.learninstruc.2012.06.001

Finney, S.J. & DiStefano, C. (2006). Non-normal and Categorical data in structural equation modeling. Greenwich, Connecticut: Information Age Publishing.

G. Hawkins, P., & Woollons, D. . (1998). Failure modes and effects analysis of complex engineering systems using functional models. *Artificial Intelligence in Engineering, 12*(4), 375–397. doi:10.1016/S0954-1810(97)10011-5

Hammond, J., Lafayette, W., Koubek, R. J., & Harvey, C. M. (2001). Distributed Collaboration for Engineering Design : A Review and Reappraisal. *Human Factors and Ergonomics in Manufacturing*, *11*(1), 35–52.

Hilts, S., Johnson, K. and Turoff, M. (1986). Experiments in group decision-making: communication process and outcome in face-to-face versus computerized conferences. *Human. Communication. Research.*, (pp. 13, 225 – 52).

Hisarciklilar, O., & Boujut, J.-F. (2009). An annotation model to reduce ambiguity in design communication. *Research in Engineering Design*, *20*(3), 171–184. doi:10.1007/s00163-009-0073-6

Initiatives-NISSAN, Quality. (2009). Retrieved from http://www.nissan-global.com/EN/QUALITY/PRODUCTS/

Mach, P., & Duraj, A. (2008). Failure mode and effects analysis of a process of reflow lead-free soldering. *2008 31st International Spring Seminar on Electronics Technology*, *4*(08), 394–398. doi:10.1109/ISSE.2008.5276636

Mohan, K., Jain, R., & Ramesh, B. (2007). Knowledge networking to support medical new product development. *Decision Support Systems*, *43*(4), 1255–1273. doi:10.1016/j.dss.2006.02.005

Ostergaard, K. J., Wetmore III, W. R., Divekar, A., Vitali, H., & Summers, J. D. (2005). An experimental methodology for investigating communication in collaborative design review meetings. *Co-Design*, *1*(3), 169–185. doi:10.1080/15710880500298520

Pahl, G., Beitz, W., Wallace, K., & Lucienne Blessing. (2007). *Engineering Design: A Systematic Approach* (3rd ed.). London: Springer-Verlag London Limited.

Peace, S. (1993). *Taguchi Methods*. Addison Wesley.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd ed.). Orlando, FL: Harcourt Brace.

Pfeiffer, T. (2002). *Quality Management: Strategies, Methods, Techniques* (3rd ed.). München, Hanser Fachbuchverlag.

Pugh, S. (1991). *Integrated Methods for Successful Product Engineering*. Addison-Wesley: London.

Sater-Black, K. and Iverson, N. (1994). How to conduct a design review, (116), 89–93.

Schmidt, C., Dart, P., Johnston, L., Sterling, L., & Thorne, P. (1999). Disincentives for communicating risk : a risk paradox. *Information and Software Technology*, *41*(February 1997), 403–411.

Schmidt, R. (2010). Risk assessment in research projects. *2010 Proceedings - Annual Reliability and Maintainability Symposium (RAMS)*, 1–6. doi:10.1109/RAMS.2010.5448039

Schmidt, R., & Spindler, C. (2012). Failure assessment and HALT test of electrical converters. *2012 Proceedings Annual Reliability and Maintainability Symposium*, 1–6. doi:10.1109/RAMS.2012.6175439

Schraw, G., Potenza, M. T., & Nebelsick-Gullet, L. (1993). Constraints on the calibration of performance. *Contemporary Educational Psychology*.

Shermer, M. (2002). *The Skeptic Encyclopedia of Pseudoscience* (p. 455).

Stamatis, D. (2003). *Failure Mode Effect Analysis: Fmea from Theory to Execution* (Second., pp. 129–153). Milwaukee 53203: American Society for Quality, Quality Press.

Teng, S.-H. (1996). Failure Mode and Effect Analysis: An Integrated Approach for Product Design and Process Control. *International Journal of Quality and Reliability*, *14*(5), 8.

Teoh, P. C., & Case, K. (2005). An evaluation of failure modes and effects analysis generation method for conceptual design. *International Journal of Computer Integrated Manufacturing*, *18*(4), 279–293.

Thorndike, E. L., & Gates, A. I. (1929). *Elementary principles of education*. New York: The Macmillan Company.

Ullman, D. G., & Spiegel, B. (2006). Trade Studies with Uncertain Information. *Sixteenth Annual International Symposium of the International Council On Systems Engineering (INCOSE)*.

Wetmore III, W. R., & Summers, J. (2003). Group decision making: friend or foe? *Engineering Management Conference, 2003. IEMC'03. Managing Technologically Driven Organizations: The Human Side of Innovation and Change* (pp. 405–409).

Wetmore III, W. R., Summers, J. D., & Greenstein, J. S. (2010). Experimental study of influence of group familiarity and information sharing on design review effectiveness. *Journal of Engineering Design*, *21*(1), 111–126.

Wetmore, W. (2004). *Proper Review Selection Matrix*. Clemson University.

Xu, K., Tang, L. ., Xie, M., Ho, S. ., & Zhu, M. . (2001). Fuzzy assessment of FMEA for engine systems. *Fuzzy assessment of FMEA for engine systems*, *75*(1), 17–29. doi:10.1016/S0951-8320(01)00101-6

Yasenchak, L. (2000). FMEA can head off potential problems before they occur. *MODERN PLASTICS*, *77*(5), 80.