DOWNLOAD PDF OR TESTING THEIR INTERACTION IN A DIFFERENT SETTING.

Chapter 1 : Experimental Research: Control, Designs, Internal and External Validity

Replication by different investigators, of different behaviors, or in different settings. Clinical Replication The development of a treatment package, composed of two or more interventions that have been found to be effective individually.

Finding Interactions The concept of an interaction can be a difficult one for students new to the field of psychology research, yet interactions are an often-occurring and important aspect of behavioral science. The following lesson will introduce the concept of a statistical interaction, provide examples of interactions, and show you how to detect an interaction. What is an interaction? When two or more independent variables are involved in a research design, there is more to consider than simply the "main effect" of each of the independent variables also termed "factors". That is, the effect of one independent variable on the dependent variable of interest may not be the same at all levels of the other independent variable. Another way to put this is that the effect of one independent variable may depend on the level of the other independent variable. In order to find an interaction, you must have a factorial design, in which the two or more independent variables are "crossed" with one another so that there are observations at every combination of levels of the two independent variables. For example, if you were interested in the effects of practice and stress level on memory task performance, you might decide to employ a factorial design. You manipulate practice by having participants read a list of words either once or five times. You also manipulate stress level by having two conditions: Your dependent variable is the number of words recalled from the word list. In this design, you would need to have participants in each of the four cells of the design: Now, if the two factors in the study practice and stress interact, this means that the effect of one factor depends on the level of the other factor. The table above indicates the cell means, as well as the marginal means and the grand mean, for the study. For example, the mean number of words recalled under the low stress, one practice condition is 8. This is a cell mean. However, the mean number of words recalled under all low stress conditions regardless of practice is This is a marginal mean. So, do we have evidence of an interaction in this study? One way to answer this question is to begin by describing the main effects: It appears that there may be a main effect of stress. High stress conditions result in recall of fewer words than low stress conditions. It also appears that there is a main effect of practice: However, the effect of the practice variable depends on the level of stress and vice versa: Therefore, we have evidence of an interaction in this study. Of course, you will need to carry out the appropriate statistical test before you can conclude that your evidence is strong enough to support the claim that there is an interaction in the population. You may want to know if there are other ways to detect this interaction besides examining the cell means. Using graphs to detect possible interactions Visually inspecting the data using bar graphs or line graphs is another way of looking for evidence of an interaction. Each of the graphs below Plots depicts a different situation with regard to the main effects of the two independent variables and their interaction. You can visualize the main effects and interaction effects if there are any in both the line graphs as drawn and in the bar graphs, which are made visible by hovering over the "View as bar graph" button.

Chapter 2 : Laboratory Testing for Prescription Opioids

System testing is the testing to ensure that by putting the software in different environments (e.g., Operating Systems) it still works. System testing is done with full system implementation and environment.

Overview[edit] Although testing can determine the correctness of software under the assumption of some specific hypotheses see hierarchy of testing difficulty below , testing cannot identify all the defects within software. These oracles may include but are not limited to specifications, contracts , [3] comparable products, past versions of the same product, inferences about intended or expected purpose, user or customer expectations, relevant standards, applicable laws, or other criteria. A primary purpose of testing is to detect software failures so that defects may be discovered and corrected. Testing cannot establish that a product functions properly under all conditions, but only that it does not function properly under specific conditions. In the current culture of software development, a testing organization may be separate from the development team. There are various roles for testing team members. Information derived from software testing may be used to correct the process by which software is developed. For example, the audience for video game software is completely different from banking software. Therefore, when an organization develops or otherwise invests in a software product, it can assess whether the software product will be acceptable to its end users, its target audience, its purchasers and other stakeholders. Software testing aids the process of attempting to make this assessment. Defects and failures[edit] Not all software defects are caused by coding errors. One common source of expensive defects is requirement gaps, e. Software faults occur through the following processes. A programmer makes an error mistake , which results in a defect fault, bug in the software source code. If this defect is executed, in certain situations the system will produce wrong results, causing a failure. For example, defects in dead code will never result in failures. A defect can turn into a failure when the environment is changed. Examples of these changes in environment include the software being run on a new computer hardware platform, alterations in source data , or interacting with different software. Input combinations and preconditions[edit] A fundamental problem with software testing is that testing under all combinations of inputs and preconditions initial state is not feasible, even with a simple product. More significantly, non-functional dimensions of quality how it is supposed to be versus what it is supposed to do â€" usability , scalability , performance , compatibility , reliability â€"can be highly subjective; something that constitutes sufficient value to one person may be intolerable to another. Combinatorial test design enables users to get greater test coverage with fewer tests. Whether they are looking for speed or test depth, they can use combinatorial test design methods to build structured variation into their test cases. More than a third of this cost could be avoided, if better software testing was performed. Until the s, the term "software tester" was used generally, but later it was also seen as a separate profession. Regarding the periods and the different goals in software testing, [11] different roles have been established, such as test manager, test lead, test analyst, test designer, tester, automation developer, and test administrator. Software testing can also be performed by non-dedicated software testers. Testing approach[edit] Static vs. Reviews , walkthroughs , or inspections are referred to as static testing, whereas actually executing programmed code with a given set of test cases is referred to as dynamic testing. Dynamic testing takes place when the program itself is run. These two approaches are used to describe the point of view that the tester takes when designing test cases. A hybrid approach called grey-box testing may also be applied to software testing methodology. White-box testing White-box testing also known as clear box testing, glass box testing, transparent box testing, and structural testing verifies the internal structures or workings of a program, as opposed to the functionality exposed to the end-user. In white-box testing, an internal perspective of the system the source code , as well as programming skills, are used to design test cases. The tester chooses inputs to exercise paths through the code and determine the appropriate outputs. While white-box testing can be applied at the unit , integration , and system levels of the software testing process, it is usually done at the unit level. Though this method of test design can uncover

many errors or problems, it might not detect unimplemented parts of the specification or missing requirements. Techniques used in white-box testing include: This allows the software team to examine parts of a system that are rarely tested and ensures that the most important function points have been tested. This is helpful in ensuring correct functionality, but not sufficient since the same code may process different inputs correctly or incorrectly. Pseudo-tested functions and methods are those that are covered but not specified it is possible to remove their body without breaking any test case. Black-box testing Black box diagram Black-box testing also known as functional testing treats the software as a "black box", examining functionality without any knowledge of internal implementation, without seeing the source code. The testers are only aware of what the software is supposed to do, not how it does it. Test cases are built around specifications and requirements, i. It uses external descriptions of the software, including specifications, requirements, and designs to derive test cases. These tests can be functional or non-functional , though usually functional. Specification-based testing may be necessary to assure correct functionality, but it is insufficient to guard against complex or high-risk situations. Whatever biases the programmers may have had, the tester likely has a different set and may emphasize different areas of functionality. On the other hand, black-box testing has been said to be "like a walk in a dark labyrinth without a flashlight. This method of test can be applied to all levels of software testing: Component interface testing Component interface testing is a variation of black-box testing , with the focus on the data values beyond just the related actions of a subsystem component. One option for interface testing is to keep a separate log file of data items being passed, often with a timestamp logged to allow analysis of thousands of cases of data passed between units for days or weeks. Tests can include checking the handling of some extreme data values while other interface variables are passed as normal values. Visual testing[edit] The aim of visual testing is to provide developers with the ability to examine what was happening at the point of software failure by presenting the data in such a way that the developer can easily find the information she or he requires, and the information is expressed clearly. Visual testing, therefore, requires the recording of the entire test process â€” capturing everything that occurs on the test system in video format. Output videos are supplemented by real-time tester input via picture-in-a-picture webcam and audio commentary from microphones. Visual testing provides a number of advantages. The quality of communication is increased drastically because testers can show the problem and the events leading up to it to the developer as opposed to just describing it and the need to replicate test failures will cease to exist in many cases. The developer will have all the evidence he or she requires of a test failure and can instead focus on the cause of the fault and how it should be fixed. Ad hoc testing and exploratory testing are important methodologies for checking software integrity, because they require less preparation time to implement, while the important bugs can be found quickly.

Chapter 3 : Software testing - Wikipedia

Stack Exchange network consists of Q&A communities including Stack Overflow, the largest, most trusted online community for developers to learn, share their knowledge, and build their careers.

The Sociological Perspective This section of the course introduces students to the discipline of sociology, focusing on its history, the questions and scientific methods that characterize it as a field, and what distinguishes it from other social science disciplines. Included in this definition is the ongoing evolution of sociology as a discipline that is both basic science and applied science. Important in this perspective are the elements of sociological practice and possible careers in sociology at all levels of academic preparation. The first two units of the course introduce students to the dynamic interplay between theory and the logic of the scientific method in sociology. Learners will become aware of the core theoretical perspectives and the process of developing theory. They will recognize that sociology is a science: The history of sociology is grounded in social and ideological changes in Western Europe and America, specifically the Enlightenment and American pragmatism. Contributions of classical sociological theorists such as Durkheim, Marx, and Weber are examined in combination with major scholars prominent in the emergence of American sociology. Sociological theory attempts to explain in a coherent manner the varieties of societal organization and of social behaviors. Students should understand that though it is posed at an abstract level, sociological theory is continually being refined as it is made to confront empirical reality. Students should become familiar with the major sociological approaches --functionalism, conflict theory, symbolic interactionism, exchange theory, and feminist theory -- to the explanation of social life. With functionalism Durkheim, Parsons , students should be aware of the analogy of society to an organism, the assumption of consensus that underlies social life, and ways that society organizes itself to sanction deviance so that it may return to equilibrium. Students should also be aware of the criticisms of functionalism regarding its difficulty in dealing with social change. Conflict theory Marx, Weber introduces students to the notion that societal stability may come from stable power relations rather than from an underlying consensus. Students should become aware of the multiplicity of conflicting interests in society as well as how changes in resources may, among other factors, lead to major social change. The difficulty of conflict theory in predicting precisely where the fissures in a given society are and when they may erupt is a recurring criticism. An inductive, qualitative approach to the understanding of individual and group interaction in a variety of contexts is the common orientation of symbolic interactionists. Exchange theory Blau, Homans, Coleman brings issues of rational choice to the fore. Students should understand the ways in which relationships of trust and power may develop as people pursue their self-interest. The degree to which exchange theory is relevant largely to interactions among individuals rather than groups and is contextually based in the larger culture should be understood. Feminist theory Gilman, Rossi, Millett focuses on the ways that gender systems structure our daily interactions as well as larger systems of power in society. Many feminist theorists focus not only on how patriarchal societies are set up in ways that disadvantage women but on how the effects of patriarchy articulate with other systems of domination, such as class- and race-based domination. From theories of sexual politics to sociobiology to economic and materialist approaches, feminist theory provides a variety of perspectives on relations of power in society. Feminist theories differ radically in how they incorporate other approaches to the study of social life. Research Methods Learners will connect the use and construction of theory with the application of diverse research methods to answer sociological questions. Over the years, philosophers, religious leaders, journalists, and many others have speculated about human society. Students will learn how sociology differs from these other enterprises because sociology applies relevant theories and scientific methods to the study of society. The methods are not pre-determined; they depend upon the question being asked. Sometimes the endeavor is exploratory; sometimes it is to test a specific theoretical proposition; it is always systematic. Students will learn how the theory-method process develops and uses a strategy that requires stating a clear question or hypothesis,

developing data to address the question or test the hypothesis, and then judging whether the question is answered or the hypothesis is supported. They will learn further that a scientific approach requires that the methods be stated clearly so that other sociologists might repeat the study to confirm the results. Coverage includes both qualitative and quantitative research, basic and applied research contexts as well as review of different methodologies, including survey research, interviewing, participant observation, content analysis, historical and comparative research. Basic concepts of statistical analysis are also included, along with discussion of probability and measurement. In addition, the course will examine the questions of ethics in research and the role of values in sociological analysis. The scientific method operates in an ethical context. As such, it does not permit the sociologist to conceal or ignore information that fails to support the hypothesis. It also requires that sociological researchers safeguard the human subjects who are a part of their research. Also included is the use of the internet in research, with a focus on judging the reliability and validity of information found on the internet. Students will learn how hypotheses are formulated using concepts and relevant sociological theory. To put this hypothesis more concretely, the rate of juvenile delinquency will be higher in low-status neighborhoods and it will be lower in high-status neighborhoods. The rate of juvenile delinquency can be measured from publicly available information on juvenile convictions and census data on the number of teenagers in a neighborhood. The rate of juvenile delinquency is a variable, because its value will differ from place to place. So, for example, a sociologist might measure social class or status with the variables annual income and occupational prestige. In testing this hypothesis, the sociologist is expecting that neighborhoods with more high income and high occupational prestige residents will have lower juvenile delinquency rates, while neighborhoods with more low income and low occupational prestige residents will have higher rates. Hypotheses are often tested using a sample of the population of interest. Students will learn how sociologists have developed careful techniques for drawing mathematically accurate samples of the population. Administering a questionnaire to a sample of the population is much less expensive and more accurate than trying to question everyone within a population. In developing explanations, students will learn how sociologists are careful to distinguish the types of variables they are investigating. In general, a dependent variable is the variable being studied. In the example in the preceding paragraph, the dependent variable is juvenile delinquency rate, and the independent variables are family income and the occupational prestige of workers in the family. Notice, however, that high parental income may be associated with a low juvenile delinquency rate, but it does not necessarily cause a low juvenile delinquency rate. Instead, the relationship may be mediated in various ways. For example, wealthier parents may be able to provide more activities for their teens, or they may be able to hire better lawyers if their teens do get into trouble. Drawing on theoretical foundations, students will learn that to assess a causal relationship between variables, it is necessary 1 to establish the time order of the variables with the independent variable coming before the dependent variable , 2 to establish that the variables are correlated, and 3 to rule out any competing hypotheses. Suppose, for example, that a researcher finds that ice cream consumption is inversely related to juvenile delinquency rates. This finding does not prove that ice cream prevents juvenile delinquency. Instead, this hypothesis may be misspecified because the wrong independent variable has been named. Perhaps parents with higher income can buy more ice cream, so that higher income co-varies positively with more ice cream and also co-varies negatively with the juvenile delinquency rate. Students will learn that sociologists collect their data through a number of research methods. One of the most common is the social survey, in which a sample of people respond to a questionnaire that is administered on paper, in a personal interview, by telephone, or over the internet. Sociologists may also engage in a participant observation, in which they become part of a group they seek to understand. Some sociologists, like psychologists, conduct experiments, while others rely principally on historical or archival data to test their hypotheses. The choice of data collection methods depends upon the kind of data that are needed to test a hypothesis. Some hypotheses may be tested through multiple methods. Students will learn how sociologists tabulate their data using statistical methods, some of which are highly sophisticated. It is common to report measures of central tendency for each variable, for example, the mean or

median values. It is also common to report a measure of the spread from the mean, such as the standard deviation or interquartile range. A measure of association such as chi-square can show if the relationship of two variables might have happened by chance or if it is a significant relationship; it is also possible to calculate the strength of an association through the use of a correlation coefficient. Students of sociology will learn how to assess the adequacy of research reported in newspapers, websites, and other places. In general, it is important to be able to tell how the research was done, whether competing hypotheses were adequately examined, and whether the appropriate variables were controlled. Studies that contain little information about how the data were collected and analyzed â€" in particular, studies that cannot be done again by another researcher â€" are suspect. Culture This section of the course introduces students to the concept of culture as the realm of socially constructed and, to varying extents, shared ideas, understandings, mental models, modes of categorization, values, speech forms, and traditions. Culture-making is a fundamental human capacity: Cultural phenomena can be observed at any level of analysis, from small groups of short duration to large-scale national societies to the world system as a whole. Although many people, including high school students, may see culture as an unyielding part of their environment, at any scale it is a human creation, and sociologists understand it as such. This section begins by looking at culture in small groups, with examples from classic social psychological experiments of how group pressures can influence and thus shape culturally such phenomena as beliefs about the natural world for example, assessments of the length of a line or deeply held moral values for example, against inflicting pain. Small groups both produce culture as in minimal-group experiments, where randomly assigned groups create elaborate beliefs about themselves and others based on trivial cues and reflect it as in research experiments demonstrating how cultural stereotypes shape the interpretation of behavior of members of a task group. Examples are considered that show how culture arises from situational contexts, how it changes, and how it influences human behavior. The next part of this section describes various elements of culture that sociologists study. Some elements of culture are so deeply institutionalized and so much a part of elementary socialization that they shape the very ways in which people think: Sociologists inherited the notion of culture from anthropologists who studied relatively small-scale, structurally simple, societies where most people shared similar beliefs and participated in the same rituals. More recently, sociologists who study culture have focused on the ways in which group and national cultures may be varied and inconsistent. This change reflects some developments in scholarship: The change also reflects developments in the real world, especially multiculturalism and globalization. With respect to multiculturalism, a dramatic increase in transnational migration has men and women raised in many places and in many cultural traditions living side by side throughout the world. This section of the course enables the instructor to explore with students the ways in which different aspects of culture serve, at different times and in different ways, as sources of cohesion, as springboards for innovation or creativity, or as bases of social conflict. Socialization Socialization is learning to become a member of the groups and society in which one lives, and is one way that societies continue through time. The course considers such key questions as: Who or what are the primary agents of socialization? What happens when infants receive very little human contact? When does socialization occur, and how does it differ at various stages of the life cycle? What role do "rites of passage" play in transitions through the life course? How do sociologists analyze the contents and context of socialization? How do various sociological perspectives illuminate socialization processes? What do sociologists mean by resocialization and desocialization? Social Organization and Social Networks Human societies routinely accomplish what, when one thinks about it, are remarkable feats of coordination: The study of social organization provides the tools for understanding the range of forms and processes that enable people to accomplish such routine miracles of social choreography. This section begins with some basic ideas. Because many roles fit with other roles into complementary relationships, they constitute basic building blocks of larger collectivities.

Chapter 4 : Interaction (statistics) - Wikipedia

Compatibility testing ensures that software can run on a different configuration, different database, different browsers and their versions. Compatibility testing is performed by the testing team. Compatibility testing is performed by the testing team.

Beta Testing Unit Testing Unit testing is the testing of an individual unit or group of related units. It falls under the class of white box testing. Integration Testing Integration testing is testing in which a group of components are combined to produce output. Also, the interaction between software and hardware is tested in integration testing if software and hardware components have any relation. It may fall under both white box testing and black box testing. Functional Testing Functional testing is the testing to ensure that the specified functionality required in the system requirements works. It falls under the class of black box testing. System Testing System testing is the testing to ensure that by putting the software in different environments e. System testing is done with full system implementation and environment. Stress Testing Stress testing is the testing to evaluate how system behaves under unfavorable conditions. Testing is conducted at beyond limits of the specifications. Performance Testing Performance testing is the testing to assess the speed and effectiveness of the system and to make sure it is generating results within a specified time as in performance requirements. Usability Testing Usability testing is performed to the perspective of the client, to evaluate how the GUI is user-friendly? How easily can the client learn? After learning how to use, how proficiently can the client perform? How pleasing is it to use its design? This falls under the class of black box testing. Acceptance Testing Acceptance testing is often done by the customer to ensure that the delivered product meets the requirements and works as the customer expected. Regression Testing Regression testing is the testing after modification of a system, component, or a group of related units to ensure that the modification is working correctly and is not damaging or imposing other modules to produce unexpected results. Beta Testing Beta testing is the testing which is done by end users, a team outside development, or publicly releasing full pre-version of the product which is known as beta version. The aim of beta testing is to cover unexpected errors.

Chapter 5 : Interaction | Real Statistics Using Excel

The negative effect of the interaction is most easily seen when the pressure is set to 50 psi and Temperature is set to degrees. Keeping the temperature at degrees will avoid the negative effect of the interaction and help ensure a strong glue bond.

External Validity The more the results obtained in an experiment generalize to other groups or situations outside of the experimental setting, the more externally valid the experimental results. Even if the results of an experiment are an accurate gauge of what happened during that experiment, do they really tell us anything about life in the wilds of society? Thus, external validity focuses on generalizability of findings from the sample to the target population, as well as other populations. Stated differently, external validity is the degree to which results are generalizable, or applicable, to groups and environments outside the research setting. For example, if we performed an experiment on cooperative learning and lecture based instruction among fourth grade students in Statesboro, what are the chances that our results would generalize to fourth grade students in New York, California, or Florida? The better are results hold for students elsewhere, the greater the external validity of our study. High internal validity implies that the researcher has tight control on the experiment, and that the setting of the experiment may even take on an artificial flavor due to the manipulated and controlled environment. Under such conditions, it is often difficult to generalize back to real life situations. In research with human subjects, both internal and external validity must be well balanced. Note that, generally, the more control a researcher has in the experiment the more internal validity , the less likely the results will generalize to outside situations low external validity. One cannot, however, downplay the importance of internal validity i. Threats to Internal Validity A threat to the internal validity of an experiment refers to anything that can occur during an experiment that makes it difficult or impossible for the researcher to say that the experimental IV caused the changes on the DV and not something else. There are a list of common or standard threats. I will provide a brief description of each, but please read the supplemental material and the text. Here is a link that discusses internal validity and threats to internal validity: The occurrence of events that are not part of the experiment, or that occur outside of the experiment, which impact, nevertheless, upon the DV of interest. This is something the researcher cannot control, but will surely affect the outcome of the study. This is called a historical threat. History does not refer to past events but rather to things that occur during the experiment. Changes in subjects over time mental and physical changes that are not part of the experiment. This is a threat since these changes may have some affect on the variation of the DV. For example, in the study of a special instructional strategy to increase reading vocabulary from first to sixth grade, we would expect reading vocabulary to increase from first to sixth grade naturally, so it will be difficult to determine whether a treatment or simple maturation growth to cause this change in vocabulary. Only the use of a control group would help to disentangle these effects. For short term maturation to be a problem, prolonged testing over a given day could cause study participants to grow mentally tired and therefore performed poorly toward the end of the day. Such reduced performance is not due to a weaker treatment, but due to tiredness with reflects mental and physical maturation over a short term. So maturation does not refer strictly to long term changes but also to short term changes. The key to assessing maturation effects is to use a study design that includes a control group; a control group allows the research to control and assess natural maturation effects relative to the treatment group. Subjects who take a pretest in an experiment may learn things from the testing experience that will improve their posttest performance on the dependent variable. If the pretest and posttest are identical or similar, subjects may demonstrate improved scores on the posttest due to having taken the pretest. Such improvement confounds the effect of the treatment. If the instrument is invalid, then one cannot trust scores obtained from the instrument, therefore observed variation on the DV cannot be directly attributable to the treatment. Another form of instrumentation results when the calibration of an instrument changes over time, or when the way scorers assign scores changes over time. For example, suppose one is interested learning

whether a particular counselor training activity changes the way counselors interact with students in a small group counseling situation. To assess possible change pre and post training, a group of counselors are watched and rated by one trained observed prior to exposure to the counselor training session. After concluding the training session, which occurred over a period of weeks, a second trained observed watches the counselors again and rates their interactions. If there are differences in the way the two trained raters score what they see, it is possible that differences in scores for both pre- and post observations could result, and these differences may not be the result of the counselor training activity, but, instead, could be due simply to differences in the way the observers score the counselors. The possibility that results of the experiment are due to a tendency for groups, selected on the basis of extreme scores, to regress toward a more average score on subsequent measurements, regardless of the effects of the experimental treatment. A group selected because of unusually low or high performance will, on average, score closer to the mean on subsequent testing, regardless of what transpires in the experiment. When intact or already formed groups are used, they may differ on some important characteristic, or when groups are selected or formed differently. The threat here is that this initial difference will result in differences on the posttest or outcome, which will confuse the study. In short, anytime already formed groups groups that are not randomly formed by the researcher are used, we must expect the groups to differ on important characteristics, or, if people are selected in different manner for different groups, then the groups are likely to differ. Subjects simply drop out of the study. This is a threat since one will not know whether these subjects represent some important difference. Threats to external validity include: Pretesting may sensitize the experimental subjects to the experimental factor so the results obtained can be generalized only to other pretested groups. In some studies a pretest may interact with the treatment and affect the results of the dependent measure; thus, one cannot generalize the results to studies in which no pretest was administered. Subjects receive more than one treatment and the researcher has no idea which treatment caused the results. Generalizability limited due to the interaction between the subjects selection and treatment, i. For example, research shows that more intelligent children tend to benefit less from cooperative learning than do less intelligent children. This illustrates selection-treatment interaction--the treatment cooperative learning works differently with different groups. The more specialized or specific the experiment in terms of the subjects, measuring instruments, outcomes, etc. This simply means that if an experiment is narrowly defined e. The more broad the characteristics of those things used in the study, the better the generalizability to other settings or people. The behavior of the subjects may be unintentionally influenced by certain characteristics or behaviors of the experimenter. In short, the researcher could bias the study in some way. Subjects may react according to their knowledge of the experiment. If subjects know they are in an experiment, then they may act differently than had they not known they were part of the experiment i. The very fact that a person is selected to participate in an experiment often motivates him to greater efforts, so the results are not applicable to other people exposed to the same experimental factor in a non-experimental setting. Supplemental Reading 1 Brenda R. Motheral writes an article that covers basic of quantitative research. In that article, she covers internal and external validity and provides examples. Please have a read of it at: Unfortunately, many times the tendency to embark on a research endeavor lacks the necessary foresight in constructing the design. Research designs are pervious to many different types of threats to their internal and external validity. In the traditions of Campbell and Stanley, and Cook and Campbell, this paper will elucidate some of the more common types of research designs, along with the coexistent threats to validity. Further, an example of each type of design has been given from the counseling literature for the interested reader to paruse and help make the concepts concrete. A Primer on Experimental and Quasi-Experimental Design Poem by Skyler Huck, University of Tennessee Upon its inception as an experimental science, psychology has utilized the scientific method found in the physical sciences. Since that time, the scientific method has been applied to various psychological constructs, e. Since the inauguration of counseling psychology in Whitley, it was mainly an applied psychology. At this juncture in time, the application of psychology was only beginning to gain respect from the intact group of psychologists who considered themselves as "pure"--that is they engaged in experimental

psychology Benjamin, In light of this zeitgeist and the identity struggles within the division, it stands to reason that counseling psychology places scientific inquiry through the rigor of the scientific method, as a core function. In the past 20 years, there has been growing dissension in the ranks of counseling psychology researchers regarding the way in which research focusing on the philosophy of science and counseling psychology is being conducted. Indeed, the results of a national survey of counseling psychologists indicate that If this is the case, how is it that so many counseling psychologists and counseling psychology students are producing "bad science"? These errors are most likely representative and indicative of the more common mistakes found in research presently. With this backdrop and apparent need for remediation, the present paper presents a "primer on experimental designs," with the specific goal of review , and the more comprehensive intention that through "better and improved science", counseling psychology will continue to solidify its place as an indisputably separate and viable field. In the traditions of Campbell and Stanley and Cook and Campbell , a review of experimental and quasi-experimental designs and how threats to validity impacts counseling research will be presented, employing examples from the current counseling literature. The Validity of Experimental Designs Internal validity Internal validity is one important type of research validity. The term "internal validity" refers to the extent that extraneous variables error variance in an experiment are accounted for. It is paramount to the researcher that model specification error variance as distinct from measurement and sampling error variance is controlled because if not, the researcher can not emphatically conclude that the observed outcome is due to the independent variable s Parker, Campbell and Stanley stated that "internal validity is the basic minimum without which any experiment is uninterpretable" p. There are eight major threats to internal validity: In later writings, Cook and Campbell identify an additional threat to internal validity. This is ambiguity about the direction of casual influence when all other plausible third-variable explanations have been ruled out of the A-B relationship, but it remains unclear as to whether A causes B, or B causes A. External Validity This construct asks the question of generalizability. Which populations, settings, treatment variables and measurement variables can these results be generalized to? Generalizing across persons requires research samples to be representative of the population of interest. Generalizing across times and settings usually necessitates systematically administering the experimental procedure at different times and different settings Parker, The inability to obtain samples that are representative of the populations from which they came, especially if studied in various settings, and at different times, results in the inability to generalize beyond the persons, time, and setting of the original study. Tests that do meet the representativeness criteria are, in essence, tests of statistical interaction. For example, if there is an interaction between a therapeutic treatment and ethnicity, then it can not be decisively stated that the treatment holds true across different ethnicities. The statistical interaction threats to external validity outlined by Cook and Campbell are as follows: Interaction of selection and treatment as in the previous example dealing with ethnicity ; and interaction of setting and treatment e. The last interaction is between history and treatment. In this case, the question involves to which period of the past or future can the results obtained be generalized. For example, the majority of experiments take place on university campuses, with undergraduate university students as participants. Parker reviewed and synthesized the Campbell and Stanley and the Cook and Campbell work and explicated two additional threats to external validity: The interaction of treatments with treatments, which refers to the administration of multiple treatments administered to the same participants, e. In the external validity case, the pretest may increase or decrease the participants responsiveness or sensitivity to the treatment.

DOWNLOAD PDF OR TESTING THEIR INTERACTION IN A DIFFERENT SETTING.

Chapter 6 : Ecological interactions (article) | Ecology | Khan Academy

An interaction variable or interaction feature is a variable constructed from an original set of variables to try to represent either all of the interaction present or some part of it. In exploratory statistical analyses it is common to use products of original variables as the basis of testing whether interaction is present with the.

Threats to validity of Research Design Chong-ho Yu The books by Campbell and Stanley , Cook and Campbell , and Shadish, Cook, and Campbell, are considered seminal works in the field of experimental design. The following write-up is based upon their books with insertion of my examples and updated information. Problem and Background Experimental method and essay-writing Campbell and Stanley point out that adherence to experimentation dominated the field of education through the s Thorndike era but that this gave way to great pessimism and rejection by the late s. However, it should be noted that a departure from experimentation to essay writing Thorndike to Gestalt Psychology occurred most often by people already adept at the experimental tradition. Therefore, we must be aware of the past so that we avoid total rejection of any method, and instead take a serious look at the effectiveness and applicability of current and past methods without making false assumptions. Replication Lack of replicability is one of the major challenges in social science research. After replicating one hundred psychological studies, Open Science Collaboration OSC found that a large portion of the replicated results were not as strong as the original reports in terms of significance p values and magnitude effect sizes. Nonetheless, the preceding problem is not surprising because usually the initial analysis tends to overfit the model to the data. Needless to say, a theory remains inconclusive when replicated results are unstable and inconsistent. Multiple experimentation is more typical of science than a one-shot experiment! Experiments really need replication and cross-validation at various times and conditions before the theory can be confirmed with confidence. In the past the only option is to replicate the same experiments over and over. Nevertheless, today the researcher is allowed to virtually repeat the study using one single sample by resampling. Specifically, many data mining software applications have the features of cross-validation and bootstrap forest. In cross-validation the data set is partitioned into many subsets and then multiple analyses are run. In each run the model is refined by previous "training" and thus the end result is considered a product of replicated experiments. In a similar vein, bootstrap forest randomly selects observations from the data and replicate the analysis many times. The conclusion is based on the convergence of these diverse results. Cumulative wisdom An interesting point made is that experiments which produce or support opposing theories against each other probably will not have clear cut outcomes. In fact, different researchers might observe something valid that represents a part of the truth. Adopting experimentation in education should not imply advocating a position incompatible with traditional wisdom. Rather, experimentation may be seen as a process of refining or enhancing this wisdom. Therefore, cumulative wisdom and scientific findings need not be opposing forces. Factors Jeopardizing Internal and External Validity Please note that validity discussed here is in the context of experimental design, not in the context of measurement. Efficacy and effectiveness In medical studies, usually efficacy studies in experimental settings are conducted to address the issue of internal validity whereas effectiveness studies in naturalistic settings the "real" world are employed to examine the external validity of the claim. Usually patients in experimentation are highly selected whereas patients in the real world are not. For example, subjects in clinical trials usually have just the illness under study. Patients who have multiple health conditions are excluded from the study because those uncontrolled variables could muddle the research results. However, in the real world it is not unusual that patients have multiple illnesses. As a result, a drug that could work well in a lab setting may fail in the real world. Thus, medical researchers must take both internal validity and external validity into account while testing the goodness of a treatment. On one hand, efficacy studies aim to answer this question: Does the treatment work in a close experimental environment? On the other hand, effectiveness studies attempt to address a different issue: Does the treatment work in the real-life situation? Interestingly enough, the US drug

approval and monitoring processes seem to compartmentalize efficacy and effectiveness. But after the drugs are on the market, it takes other agencies to monitor the effectiveness of the drugs. Contrary to the popular belief, FDA has no authority to recall unsafe drugs. Rather, FDA could suggest a voluntarily recall only. Several drugs that had been approved by FDA before were re-called from the market later e. This discrepancy between the results yielded from lab tests and the real world led to an investigation by the Institute of Medicine IOM. To close the gap between internal and external validity, the IOM committee recommended that the FDA should take proactive steps to monitor the safety of the approved drugs throughout their time on the market Ramsey, Indeed, there is a similar concept to "effectiveness" in educational research: Educational researchers realize that it is impossible for teacher to blocking all interferences by closing the door. Whether internal validity or external validity is more important has been a controversial topic in the research community. Campbell and Stanley stated that although ideally speaking a good study should be strong in both types of validity, internal validity is indispensable and essential while the question of external validity is never completely answerable. External validity is concerned with whether the same result of a given study can be observed in other situations. Like inductive inference, this question will never be conclusive. No matter how many new cases concur with the previous finding, it takes just one counter-example to weaken the external validity of the study. Cronbach is opposed to this notion. He argued that if a treatment is expected to be relevant to a broader context, the causal inference must go beyond the specific conditions. If the study lacks generalizability, then the so-called internally valid causal effect is useless to decision makers. In a similar vein, Briggs asserted that although statistical conclusion validity and internal validity together affirms a causal effect, construct validity and external validity are still necessary for generalizing a causal conclusion to other settings. Factors which jeopardize internal validity History: The economic recession is a good example. Due to the budget crisis many schools cut back resources. A treatment implemented around that period of time may be affected by a lack of supporting infrastructure. In other words, the pretest becomes a form of "treatment. It is also known as regression towards the mean. This phenomenon was first discovered by British statistician Francis Galton in the 19th century. Contrary to popular belief, Galton found that tall parents do not necessary have tall children. If the parent is extremely tall, the offspring tend to closer to the average. This pattern was re-discovered by Jewish-American psychologist Daniel Kahneman in his study about why rebuking pilots cannot explain flight performance. In the context of research design, the threat of regression towards the mean is caused by the selection of subjects on the basis of extreme scores or characteristics. If there are forty poor students in the treatment program, it is likely that they will show some improvement after the treatment. However, if the students are extremely poor and thus are unresponsive to any treatment, then it is called the floor effect. Randomization Random assignment of group membership is a counter-attack against this threat. However, when the sample size is small, randomization may lead to Simpson Paradox, which has been discussed in an earlier lesson. For example, in a Web-based instruction project entitled Eruditio, it started with subjects and only 95 of them completed the entire module. Those who stayed in the project all the way to end may be more motivated to learn and thus achieved higher performance. The hidden variable, intention to treat, might skew the result. John Henry effect and Hawthorne effect: John Henry was a worker who outperformed a machine under an experimental setting because he was aware that his performance was compared with that of a machine. The Hawthrone effect is similar to John Henery effect in the sense that the participants change their behaviors when they are aware of their role as research subjects. Between and 32 the Hawthorne Works sponsored a study to examine how lighting would influence productivity. Researchers concluded that workers improved their productivity because they were observed rather than better illumination. Hence, the Hawthorne effect is also known as the observer effect. Factors which jeopardize external validity Reactive or interaction effect of testing: Interaction effects of selection biases and the experimental variable Reactive effects of experimental arrangements: Three Experimental Designs To make things easier, the following will act as representations within particular designs:

Chapter 7 : What is software testing? What are the different types of testing? - CodeProject

regression models with several explanatory variables and their interactions. All data, results (in PDF format), syntax, and additional information are available through the Internet for free.

The null hypothesis itself does not involve the data. It is a statement about a parameter a numerical characteristic of the population. These population values might be proportions or means or differences between means or proportions or correlations or odds ratios or any other numerical summary of the population. Here are some examples. Suppose a researcher at Penn State speculates that students in the College of Arts and Architecture are more likely to be left-handed than people found in the general population. We only have one sample since we will be comparing a population proportion based on a sample value to a known population value. Are artists more likely to be left-handed than people found in the general population? Classification of student as either right-handed or left handed State Null and Alternative Hypotheses Null Hypothesis: This is a one-sided alternative hypothesis. The manufacturer is worried that the machine that fills the capsules has come out of calibration and is no longer creating capsules with the appropriate dosage. Does the data suggest that the population mean dosage of this brand is different than 50 mg? This is a two-sided alternative hypothesis. Hypotheses with Two Samples of One Categorical Variable Many people are starting to prefer vegetarian meals on a regular basis. Specifically, a researcher believes that females are more likely than males to eat vegetarian meals on a regular basis. Does the data suggest that females are more likely than males to eat vegetarian meals on a regular basis? Classification of whether or not a person eats vegetarian meals on a regular basis Explanatory Grouping Variable: Does the data suggest that, on the average, people are able to lose more weight on a low carbohydrate diet than on a low fat diet? Weight loss pounds Explanatory Grouping Variable: A case-control study of non-smoking stroke patients and controls of the same age and occupation are asked if someone in their household smokes. There are two different categorical variables Stroke patient vs control and whether the subject lives in the same household as a smoker. Living with a smoker or not is the natural explanatory variable and having a stroke or not is the natural response variable in this situation. This is a one-tailed alternative. Note that this research question might also be addressed like example Daily price change information the response variable and previous day stock purchases by non-management employees explanatory variable. These are two different measurement variables. Is the strength of this association different for family restaurants than for fine dining restaurants? There are two different measurement variables. The size of the tip would depend on the size of the bill so the amount of the bill would be the explanatory variable and the size of the tip would be the response variable.

Chapter 8 : Usability testing - Wikipedia

The set of predictors and all their implied interactions (in a "full model") may explain an impressive amount of the variance of the dependent variable Y, while none of the regression coefficients are significantly different from zero.

Interactions involving a dummy variable multiplied by a measurement variable are termed slope dummy variables, [12] because they estimate and test the difference in slopes between groups 0 and 1. Centering makes the main effects in interaction models more interpretable. The coefficient a in the equation above, for example, represents the effect of x1 when x2 equals zero. Centering can also reduce problems with multicollinearity. Interaction of education and political party affecting beliefs about climate change Regression approaches to interaction modeling are very general because they can accommodate additional predictors, and many alternative specifications or estimation strategies beyond ordinary least squares. Robust , quantile , and mixed-effects multilevel models are among the possibilities, as is generalized linear modeling encompassing a wide range of categorical, ordered, counted or otherwise limited dependent variables. Interaction of species and air temperature and their effect on body temperature[edit] Consider a study of the body temperature of different species at different air temperatures. The data are shown in the table below. The interaction plot may use either the air temperature or the species as the x axis. The second factor is represented by lines on the interaction plot. There is an interaction between the two factors air temperature and species in their effect on the response body temperature , because the effect of the air temperature depends on the species. The interaction is indicated on the plot because the lines are not parallel. In the interaction plot, the lines for the mild and moderate stroke groups are parallel, indicating that the drug has the same effect in both groups, so there is no interaction. The line for the severe stroke group is not parallel to the other lines, indicating that there is an interaction between stroke severity and drug effect on survival. The line for the severe stroke group is flat, indicating that, among these patients, there is no difference in survival between the drug and placebo treatments. In contrast, the lines for the mild and moderate stroke groups slope down to the right, indicating that, among these patients, the placebo group has lower survival than drug-treated group. Hypothesis tests for interactions[edit] Analysis of variance and regression analysis are used to test for significant interactions. Interaction of temperature and time in cookie baking[edit] Is the yield of good cookies affected by the baking temperature and time in the oven? The table shows data for 8 batches of cookies. The data show that the yield of good cookies is best when either i temperature is high and time in the oven is short, or ii temperature is low and time in the oven is long. If the cookies are left in the oven for a long time at a high temperature, there are burnt cookies and the yield is low. From the graph and the data, it is clear that the lines are not parallel, indicating that there is an interaction. Based on the interaction test and the interaction plot, it appears that the effect of time on yield depends on temperature and vice versa. Examples[edit] Real-world examples of interaction include: Interaction between adding sugar to coffee and stirring the coffee. Neither of the two individual variables has much effect on sweetness but a combination of the two does. Interaction between adding carbon to steel and quenching. Neither of the two individually has much effect on strength but a combination of the two has a dramatic effect. Interaction between smoking and inhaling asbestos fibres: Both raise lung carcinoma risk, but exposure to asbestos multiplies the cancer risk in smokers and non-smokers. Here, the joint effect of inhaling asbestos and smoking is higher than the sum of both effects. The western dietary pattern was shown to increase diabetes risk for subjects with a high "genetic risk score", but not for other subjects. For example, US surveys often find that acceptance of the reality of anthropogenic climate change rises with education among moderate or liberal survey respondents, but declines with education among the most conservative.

Chapter 9 : Understanding Interactions

Future work could help to establish the conditions under which these classroom interactions occur in different settings, and at what point these interactions serve a facilitative role in children's academic and socio-emotional development.

Abstract Opioid analgesic misuse has risen significantly over the past two decades, and these drugs now represent the most commonly abused class of prescription medications. They are a major cause of poisoning deaths in the USA exceeding heroin and cocaine. Laboratory testing plays a role in the detection of opioid misuse and the evaluation of patients with opioid intoxication. Laboratories use both immunoassay and chromatographic methods e. Testing methods for opioids originated in the workplace-testing arena and focused on detection of illicit heroin use. Analysis for a wide range of opioids is now required in the context of the prescription opioid epidemic. Testing methods have also been primarily based upon urine screening; however, methods for analyzing alternative samples such as saliva, sweat, and hair are available. Application of testing to monitor prescription opioid drug therapy is an increasingly important use of drug testing, and this area of testing introduces new interpretative challenges. In particular, drug metabolism may transform one clinically available opioid into another. The sensitivity of testing methods also varies considerably across the spectrum of opioid drugs. An understanding of opioid metabolism and method sensitivity towards different opioid drugs is therefore essential to effective use of these tests. Improved testing algorithms and more research into the effective use of drug testing in the clinical setting, particularly in pain medicine and substance abuse, are needed. Opioids, Screening, Pain medicine, Substance abuse Introduction Chronic noncancer pain represents a major health problem worldwide [1]. Although the long-term consequences of treating chronic noncancer pain with opioid drugs are unclear, prescription opioid use has risen over the past decade in an effort to enhance the treatment of chronic pain [2 , 3]. Over the past decade, misuse of prescription opioid drugs has also skyrocketed [4]. The desire to provide adequate control of pain while avoiding opioid misuse leads to a significant dilemma for the health care provider. Methods for detecting misuse are therefore critical adjuncts to prescribing these drugs. Several tools have been developed to aid in the detection opioid misuse and abuse including question-based screening tools [5 â€” 9], prescription drug monitoring programs [10], and drug testing. Urine drug testing UDT has been widely advocated as a method for identifying the misuse and abuse of opioid drugs during chronic pain and substance abuse treatment [11], and it is also frequently used in other clinical settings to screen for aberrant drug use such as in patients with behavioral disorders and other clinical signs and symptoms that suggest drug intoxication. The correct interpretation of testing is also critical as it may be the foundation to treatment decisions, criminal prosecution e. This review will focus on opioid drug testing, which continues to evolve alongside the growth in opioid drug use. In particular, the importance of testing approaches and their impact on the interpretation of results in the clinical settings will be discussed. Urine Screening for Opioids Traditional approaches to drug screening in clinical medicine have relied upon the approaches originally developed for the workplace. Although this review will focus on opioids, similar testing can be performed for a diversity of substances. In the traditional testing scheme, two tiers of testing are employed. The first tier aims to rapidly screen large numbers of specimens for opioids. The second tier of testing is generally performed using highly specific methods such as gas chromatography or liquid chromatography with mass spectrometry that serve to confirm the screening result. Since confirmatory results are not generally available the same day, and therefore less likely to influence clinical decision making, hospital laboratories performing urine drug testing solely for clinical purposes may not routinely perform this second tier of testing [13]. The initial screening step plays a pivotal role in any testing scheme. In workplace opioid testing, detection of illicit drug use is sought in a population with a generally low prevalence of opioid use [14]. In the clinical setting, the prevalence of drug exposure is much higher due to the preselection of patients for screening based upon clinical suspicion of drug exposure or prescription of the opioid drug [15]. The purpose of testing may also be different in the clinical setting where

the goal is often to detect nonuse of a prescribed opioid drug that may indicate drug diversion, which is an important contributor to the ongoing prescription opioid drug epidemic. These differences in prevalence and testing goal have an important impact on the utility of testing [16]. The effectiveness of the test can theoretically be improved in the clinical setting through changes to the concentration cutoff used to define a positive or negative test. This threshold was also developed in adult populations and may not be appropriate for children who produce less concentrated urine [17]. Lower thresholds have been advocated in the pain medicine setting [18]. The cutoff concentration has been given significantly more attention in the workplace setting. The medico-legal consequences of testing in the workplace encourage great attention towards control of false positive results. Commonly prescribed drugs such as fluoroquinolone antibiotics have been reported to produce false positive results in some opiate immunoassays [19]. While confirmatory testing can correct for these false positive immunoassay results, this comes at significant cost since the confirmatory testing is a highly labor-intensive process unlike automated immunoassay testing. It is therefore highly desirable to minimize false positive results. In the context of heroin as the most significantly abused opioid drug in the s and s, this change in cutoff concentration was predicted to significantly improve the positive predictive value i. However, the epidemiology of opioid abuse has changed, and testing guidelines will likely have to change as well to enhance detection of prescription opioids that now exceed heroin in abuse [21]. Although often misunderstood, screening tests for opioid drugs also do not detect all opioid drugs equally. Due to the limited cross-reactivity of antibodies with the diversity of opioid drugs, urine specimens containing many drugs may escape detection by opiate immunoassays. The widely prescribed and abused opioid, oxycodone, illustrates this analytical pitfall well. Due to the structural differences between oxycodone and morphine shown in Fig. A greater than 6-fold higher concentration of oxycodone is therefore required to achieve a positive screening result in many opiate immunoassays compared with morphine. Several studies have confirmed the low detection efficiency of morphine-specific opiate immunoassays for oxycodone [23 â€” 25]. Other prescription opioid drugs such as fentanyl and buprenorphine are sufficiently distinct in structure compared to morphine that these drugs show essentially no reactivity in commonly marketed morphine-specific opiate immunoassays Fig.