Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2019 Proceedings | Professional Development Workshops (PDW)

Designing Experimental Studies

Fenne große Deters
University of Potsdam, fenne.grosse.deters@uni-potsdam.de

Stefan Tams
HEC Montreal, stefan.tams@hec.ca

Allen Johnston
University of Alabama, ajohnston@cba.ua.edu

Jason Thatcher
Culverhouse College of Business, University of Alabama, jason.b.thatcher@gmail.com

Designing Experimental Studies

PDW

Fenne große Deters
University of Potsdam/Weizenbaum Institute for the Networked Society
Hardenbergstraße 32, 10623 Berlin, Germany
fenne.grosse.deters@uni-potsdam.de

Stefan Tams
HEC Montréal - École des Hautes Études commerciales
3000, chemin de la Côte Sainte Catherine, Montréal (Québec), Canada
stefan.tams@hec.ca

Allen C. Johnston
Culverhouse College of Business/University of Alabama
361 Stadium Dr, Tuscaloosa, AL 35487, USA
ajohnston@cba.ua.edu

Jason B. Thatcher
Culverhouse College of Business/University of Alabama
361 Stadium Dr, Tuscaloosa, AL 35487, USA
jbthatcher1@culverhouse.ua.edu

Abstract

In the last years, experiments became more and more widely applied - be it in academic research or A/B testing in companies. Due to their high internal validity, experiments are an important part of the methods ecosystem and researchers will benefit from integrating them into their methodological tool kit. This paper aims to summarize the most important content of the ICIS 2019 Professional Development Workshop. The workshop targets researchers with no or very basic training in experimental methods. It introduces the essentials of understanding and planning state-of-the-art experimental research and covers common pitfalls and challenges.

Keywords: Experimental Design; Research Methods; Field Experiment; Laboratory Experiment

Introduction

Experiments are an essential research method in many empirical research disciplines like physics, medicine, agriculture, and chemistry and have helped tremendously to advance these fields (Levy and Ellis 2011). Despite other research designs to study cause-effect relationships, true experiments are still widely considered the gold standard if causal conclusions are of interest (Rubin, 2008). While experiments in disciplines involving human subjects are not without challenges, they are highly popular in many neighboring fields of IS like psychology (Coolican 2017) and behavioral economics (Gupta et al. 2018) to complement other research methods with lower internal validity. In the past, experiments have been criticized for being too artificial and too removed from real life experiences (Dennis and Valacich 2001). In a discipline like IS which highly values practical relevance of results (Benbasat and Zmud 1999), this critic probably kept many researchers from using experiments in their research. However, in the last decade new research areas in IS like for example e-commerce (Burtch et al. 2017), social media (Bapna et al. 2017), cybersecurity (Wright et al. 2014), online privacy (Hui et al. 2007), mobile advertising (Sutanto et al. 2013), and online dating (Bapna et al. 2016) have opened up new venues to conduct experimental research in natural settings with dramatically increased realism and generality (Karahanna et al. 2018). Moreover, the trend to multi-study papers enables IS researchers in any research domain to combine experiments with, for example, large-scale surveys. Hence, while benefitting from the high internal validity of an experiment its potential lack of ecological validity can be complemented by adding an element of realism with other research methods (Dennis and Valacich 2001). These changes are reflected in a growing body of excellent experimental studies in IS (Adomavicius et al. 2013; Benlian et al. 2012; Cason et al. 2011; Cwiakowski et al. 2016; Fu 2011; Gregg and Walczak 2008; Hashim et al. 2017; Kamis et al. 2008; Mangalaraj et al. 2014; Rice 2012; Tams et al. 2018; Thatcher et al. 2018; Tsai et al. 2011). Moreover, many big internet-based businesses like Facebook, Amazon, Google, eBay, Groupon, LinkedIn, Netflix, Shop Direct, and Uber rely heavily on randomized experiments – termed A/B testing in this context (Deb et al. 2018; Feitelson et al. 2013; Kohavi et al. 2013; Tang et al. 2010). They demonstrate that valuable insights for business decisions can be gained from applying experimental methods nowadays. Hence, many researchers in IS might benefit from expanding their methodological tool kit by adding experiments to their repertoire of methodologies.

Experimental versus Observational studies

Quantitative researchers value internal validity, external validity (generalizability), and realism in their studies. Unfortunately, there is a trade-off: Maximizing any one of these aspects jeopardizes the other two (Dennis and Valacich 2001; Karahanna et al. 2018). Different research methods like experiments and observational studies focus on maximizing different aspects. The main asset of experimental studies – if designed and executed carefully- is their high internal validity (Karahanna et al. 2018). Researchers conduct experiments to be able to confidently draw the conclusion that the observed differences in the dependent variable (outcome) are *caused* by differences in the independent variable (Brewer and Crano 2000). The following conditions need to be met to assume causation: 1.) Changes in the independent variable need to precede changes in the dependent variable. 2.) There is a systematic relationship between variation in the independent and in the dependent variable. And 3.) rivaling explanations for the systematic association between independent and dependent variable can be ruled out (Shadish et al. 2002). Observational studies can meet criterion 2 and in the case of longitudinal studies also criterion 1, temporal order (Brewer and Crano 2000). However, in non-experimental studies alternative explanations can never be ruled out completely. It is always possible that the observed association between the independent variable and the dependent variable only exists because a third variable which is correlated with the independent variable effects the dependent variable. This so-called “Third-Variable-Problem” (Brewer and Crano 2000) is best understood when illustrated with an example: Observing a correlation between ice-cream consumption and forest fires should not lead to the conclusion that there is a causal link and hence, ice-cream should be banned. Most likely, the third variable “hot weather” causes people to eat more ice-cream and also creates dry conditions that increase the risk of forest fires (Feenstra 2015). While in this example it is obvious that correlation does not equal causation and it is straight forward to identify the third variable (Brewer and Crano 2000), for many research questions in IS this is not the case and several often unknown variables and processes could be hidden causes of the observed relationship. Experiments offer a unique opportunity to minimize the amount of rivalling explanations and provide insights into causal relationships (Shadish et al. 2002).

However, as mentioned before, the advantage of high internal validity in experimental studies often comes at the cost of lower realism and lower external validity (Brewer and Crano 2000; Karahanna et al. 2018). The classic experiment in the lab scores the highest in internal validity and the lowest in realism and external validity, whereas lab-in-the-field and field experiments sacrifice some control in favor of higher realism and external validity (Karahanna et al. 2018). Moreover, often experiments are not feasible due to practical or ethical concerns i.e., the independent variable cannot or should not be manipulated by the researcher (Shadish et al. 2002). Hence, almost always a combination of different research methods – experiments in the lab and in the field as well as observational studies- is the best choice to advance knowledge (Dennis and Valacich 2001).

Essential Elements of Experiments

In order to obtain high internal validity, true experiments are characterized by three essential elements: 1) The independent variable is purposefully manipulated or changed by the researcher and there are at least two different conditions: An experimental/treatment condition and a control condition. 2) Other extraneous variables are held constant or uncorrelated with the independent variable. And 3.) Participants have an equal chance of experiencing the different levels of the independent variable, most often assured by random assignment to the conditions (Brewer and Crano 2000; Campbell and Stanley 1966; Price et al. 2017). In the following paragraphs, these elements will be explained in detail.

Manipulation

In experimental studies, the researcher manipulates the independent variable to observe whether this results in changes in the dependent variable (Shadish et al. 2002). At minimum, true experiments have two conditions: The control condition represents the "no treatment" or neutral state (business as usual) and the experimental/treatment condition is the condition in which the level of the independent variable of interest is realized (Price et al. 2017). Differences in the dependent variable between control and experimental condition can then be attributed to the manipulation of the independent variable. A control condition is essential to assure internal validity (Campbell and Stanley 1966; Shadish et al. 2002). Changes in the dependent variable can be due to the Hawthorne effect (participants behaving differently just because they know they take part in a study) (Harrison and List 2004), maturation (a change in the dependent variable occurs naturally over time), attrition (participants dropping out of the study in a non-random way), and regression to the mean (a natural shift away from extreme scores at the pretest) to name just a few alternative causes for change. Please refer to Campbell and Stanley (1966) for a more complete list. By comparing control and experimental condition it is possible to distinguish between these effects and changes in the dependent variable that are caused by the independent variable and only occur in the experimental condition (Campbell and Stanley 1966).

For the operational definition of the independent variable in experimental research just like in observational studies, researchers need to worry about construct validity (Brewer and Crano 2000; Price et al. 2017). In addition, there are two main points that deserve particular attention in experimental research: Firstly, it is important to be aware of demand characteristics. If the manipulation of the independent variable is too obvious participants might guess the hypothesis. Because participants are often motivated to cooperate they might change their behavior or responses in accordance with what they think the researcher expects (Orne 1962). This is a serious threat to internal validity which might be alleviated by more subtle manipulations or a cover story (Durgin et al. 2009). Secondly, it is crucial to make sure that the manipulation does indeed result in changing the independent variable as intended (Shadish et al. 2002). It is important to check whether participants follow the instructions (e.g., logs of app usage), whether they paid intention to a stimuli and were exposed to the treatment (e.g., test whether they have seen an icon), and whether the manipulation has the intended effect on an internal state (e.g., are participants really bored after a "boredom manipulation"). However, manipulation checks during the experiment can act as treatments themselves or eliminate, intensify, or interact with the effects of a manipulation. While recordings of behavior are unproblematic, especially verbal manipulation checks are often better limited to extensive pilot studies before the actual experiment (Hauser et al. 2018).

Control of Extraneous Variables

Extraneous variables can introduce systematic and unsystematic variation (Price et al. 2017). Unsystematic variation i.e., variation that is unrelated to the independent variable, creates "noise" in the data and the true effect of the manipulation might be hidden (Shadish et al. 2002). Hence, researchers are well-advised to take measures to keep situation, task, and context variables constant for all participants like e.g., conducting the experiment at the same location and time as well as standardizing the exact procedure (Price et al. 2017). However, unsystematic variation reduces statistical power but does not threaten internal validity. In contrast, anything that differs *systematically* between the experimental conditions other than the levels of the independent variable should be avoided at all costs. Therefore, experience for participants in the experimental and control condition should be as similar as possible – except for the treatment itself. Otherwise conclusions about cause and effect might be false (Shadish et al. 2002).

Random Assignment

Not only context, setting, and task variables but characteristics of the participants will introduce variation (Price et al. 2017). To a certain extent it also is possible to keep these variables constant by limiting the sample to participants with certain characteristics like e.g., middle-aged employees with no prior programming experience. However, this approach dramatically decreases external validity because findings can only be generalized to a population that shares these characteristics (Price et al. 2017). Moreover, controlling all potentially influential variables is impossible. The most common solution to this problem is to randomly assign participant to the conditions (Shadish et al. 2002). Without randomization certain characteristics of participants might make it more likely that they are selected or self-select to participate in one particular condition (selection bias) (Suresh 2011). In this case differences between the groups might be due to systematic pre-treatment differences instead of the manipulation of the independent variable (Karahanna et al. 2018; Suresh 2011). Randomly assigning participants, e.g., by flipping a coin, avoids this important threat to internal validity (Price et al. 2017). It is important to note that random assignment does not guarantee that participants in both conditions are similar in each individual study. However, differences based on random processes "average out" with larger samples and multiple studies (Price et al. 2017; Shadish et al. 2002).

Especially with smaller samples, simple randomization might result in unequal group sizes (Price et al. 2017). To avoid this, block randomization can be applied (Suresh 2011). Within one block all conditions appear equally often but in a random order. Another more advanced way of random assignment is matching. If it is known that a particular characteristic of participants strongly influences the dependent variable but cannot be manipulated, researchers want to make sure that participants in both groups are very similar in regards to this characteristic (Price et al. 2017). In order to "improve chance", researchers can measure the characteristic of interest, group participants in pairs with similar scores, and then randomly assign one participant of each pair to the control and the other to the experimental condition (Campbell and Stanley 1966). Suresh (2011) provides an overview of further more specialized randomization strategies as well as resources for their implementation.

Experimental Designs

An important distinction between different experimental designs is whether participants are exclusively either in the experimental or in the control condition (between-subjects design) or whether each participants is in all experimental conditions (within-subjects designs)(Keren and Lewis 1993; Price et al. 2017). For between-subject experiments random assignment of participants to conditions is crucial (Price et al. 2017). However, within-subject designs offer a different solution to the problem of variation introduced by characteristics of the participants: Because every participant experiences all conditions, participants function as their own control condition[1] (Charness et al. 2012; Greenwald 1976). In the following two paragraphs three different simple experimental designs are discussed. More advanced designs like the Solomon four group design or factorial designs (Campbell and Stanley 1966) are outside of the scope of this paper.

Between-Subjects Designs

Two popular and relatively simple between-subjects designs are the pretest-posttest control group design and the posttest-only control group design (Campbell and Stanley 1966). In both designs, participants are randomly assigned to either the experimental or the control condition. The researcher manipulates the independent variable for participants in the experimental condition and measures the dependent variable in both conditions at the end of the experiment. In the pretest-posttest control group design, in addition the dependent variable is also measured at the beginning of the experiment before the treatment (pretest) (Campbell and Stanley 1966; Levy and Ellis 2011). In pretest-posttest control group designs potential pre-treatment differences between conditions despite randomization can be controlled. Hence, the design has a higher statistical power (Morris 2008). Moreover, if participants with higher or lower pre-treatment scores are more likely to drop-out of the experimental condition than the control condition, this selective attrition bias can be detected (Shadish et al. 2002). A drawback of a pretest is that it might increase demand characteristics (Rosnow and Suls 1970). Moreover, sometimes a pretest is impossible or too impractical. In these cases, the posttest-only control group design is a good alternative (Shadish et al. 2002).

[1] So technically, all participants have an equal chance i.e., a chance of 100%, of experiencing the different levels of the independent variable (see point 3, essential elements of experiments).

Within-Subjects Designs

In the simplest form of a within-subject design, participants experience both conditions one after the other and the dependent variable is measured after each condition (Price et al. 2017). It is possible that experiencing one condition changes how participants react to the other condition or how they score on the measure of the dependent variable if they complete it the second time (order effects) (Shadish et al. 2002). Examples of order effects are fatigue effects (participants get tired or bored), practice effects (participants get better through practice), carryover effects (the effect of one condition remains strong throughout the next condition), and sensitization effects (participants become more alert because they are clued from earlier conditions)(Greenwald 1976; Price et al. 2017). Because order effects could be misinterpreted as an effect of the manipulation, they are a serious threat to internal validity (Price et al. 2017; Shadish et al. 2002). As a remedy, participants should to be assigned randomly to experience the conditions in different orders, so-called counterbalancing (Shadish et al. 2002). While counterbalancing is straightforward with just two conditions, the number of possible orders increases quickly with several control and experimental conditions. For these cases, Price et al. (2017) describe different ways of counterbalancing (complete, partial, full) in detail. For some manipulations the same participants cannot experience all conditions in random order and counterbalancing is impossible (e.g., manipulating experience with using a certain device), or order effects might be so pronounced that the treatment effect would be hidden (Price et al. 2017). However, for many cases within-subject designs offer the advantage that characteristics of the participants are guaranteed to be equal in both conditions (Erlebacher 1977). Moreover, in comparison to between-subject designs with the same number of participants within-subject designs have a higher statistical power (Greenwald 1976; Keren and Lewis 1993).

External and Ecological Validity

A high internal validity is crucial for experimental research (Ellis and Levy 2009). However, especially in a discipline like IS with a pronounced interest in the practical application of results (Benbasat and Zmud 1999), external and ecological validity should be taken into consideration as well. External validity refers to the question of how well the results can be generalized for example, to other populations and settings (Shadish et al. 2002). And ecological validity or realism is concerned with the similarity between the experiment and real-life (Brewer and Crano 2000). As discussed before, to a certain extent internal, external, and ecological validity have conflicting demands (Dennis and Valacich 2001; Karahanna et al. 2018). For example, more homogenous samples reduce the threat to internal validity that variation on the level of the participants influences the dependent variable but they limit generalizability to other populations (Price et al. 2017). And field studies in cooperation with a company offer often less control and -if the setting is very specific- also limited external validity but can be high in ecological validity (Dipboye and Flanagan 1979; Karahanna et al. 2018). Before designing an experiment, it is helpful to clarify to which populations and to which settings the results should be generalizable and in which real-life setting the findings could be of practical relevance. This helps to guide the choice which participants should be included (Anderson et al. 2019) as well as the decision between lab, lab-in-the-field, and field experiments (Karahanna et al. 2018). Please refer to Karahanna et al. (2018) for a detailed discussion of the advantages and disadvantages of these different experimental methods (lab, field-in-the-lab, and field) in respect to internal, external, and ecological validity.

However, there are certainly also measures to improve external and ecological validity without a tradeoff (Price et al. 2017). The external validity of the treatment can often be increase (Fontenelle et al. 1985). For example, if researchers want to study the effect of the size of icons on click-rate, they should produce a set of big and small icons that also vary in respect to their color. While the size of the icon presented to a participant would be determined by the condition, the color could be randomized. A causal effect of icon size would then generalize to icons of all used colors. And striving to recruit participants that are part of the population of interest (e.g., IT administrators instead of Mturkers or students) increases realism without necessarily affecting internal validity (Bello et al. 2009).

Ethical Considerations

Ethical considerations are always mandatory for research with human subjects (Belmont Report 1979). But for experiments this is even more important because researchers willfully manipulate the independent variable with the goal of observing a change (Shadish et al. 2002). There are three hallmarks of ethical research: Informed consent, voluntary participation, and a careful assessment of risks and benefits (Karahanna et al. 2018). Participants need to be informed about the experimental procedures and agree to take part voluntarily (informed consent) (Shadish et al. 2002). Researchers should make sure that participants are able to understand what their consent implies and are capable to fully grasp potential risks of their participation (Levine et al. 2004). This aspect needs particular attention with so-called "vulnerable populations" with reduced cognitive abilities like for example children, adolescents, or people in nursing homes (Yan and Munir 2004). Unequal power relationships can also be problematic (Levine et al. 2004): For example, students or employees might find themselves in a situation where they will experience or fear to experience severe disadvantages, if they decide against participating in the experiment. Hence, there participation cannot be assumed to be voluntary.

In ethical studies, researchers strive to minimize risks and negative effects for participants and communicate unavoidable risks openly (Belmont Report 1979; Karahanna et al. 2018; Price et al. 2017). Common risks or negative effects for participants in IS research are a breach of privacy or confidentiality (e.g., if the employer learns about a participant's intention to quit) and emotional stress or discomfort experienced during the study (e.g., if social media usage elicits envy). But also withholding a potential asset from some participants (e.g., a new feature of a dating app) needs a good justification (Shadish et al. 2002). Often ethical and methodological aspects have to be considered simultaneously when designing an experiment (Karahanna et al. 2018). For example, measures to increase validity like the deception of participants in experiments by using cover stories to disguise the research question has been critically discussed (Ariely and Norton 2007; Hertwig and Ortmann 2008) just as waiving informed consent in certain settings to increase ecological validity (Grimmelmann 2015). However, there is consensus that in all cases, debriefing participants after the experiment is mandatory (Belmont Report 1979; Grimmelmann 2015). In general, we recommend to involve an independent ethics committee before conducting (experimental) research.

Data-Analytic Methods

There are many statistical methods to analyze experimental data. However, ANOVA (Analysis of Variance), ANCOVA (Analysis of Covariance), and repeated-measures ANOVA and ANCOVA are methods that have been developed particularly for experiments and are hence easy to use and well-suited for the task (Scheffe 1999). For data from between-subjects designs ANOVA can be used. If the influence of an extraneous variable on the dependent variable should be controlled a researcher can choose to conduct an ANCOVA (Rutherford 2001). Equivalent for within-subjects designs a repeated-measures ANOVA or ANCOVA are good choices. It is possible to use most statistical software packages like e.g., R, SPSS, SAS, and STATA to conduct analysis of variance or covariance.

Conclusion

Carefully designed experiments allow researchers to confidently draw causal conclusions (Campbell and Stanley 1966; Karahanna et al. 2018; Levy and Ellis 2011; Rubin 2008). Hence, they play an important role in the methods ecosystem. This paper summarizes the content of the ICIS 2019 Professional Development Workshop on designing experimental studies. While far from comprehensive, it offers interested researchers a first introduction into the topic and hopefully encourages them to venture deeper into the exciting and rewarding world of experimental research.

References

Adomavicius, G., Curley, S. P., Gupta, A., and Sanyal, P. 2013. "Impact of Information Feedback in Continuous Combinatorial Auctions: An Experimental Study of Economic Performance," *MIS Quarterly* (37:1), pp. 55-76.

Anderson, C. A., Allen, J. J., Plante, C., Quigley-McBride, A., Lovett, A., and Rokkum, J. N. 2019. "The Mturkification of Social and Personality Psychology," *Personality and Social Psychology Bulletin* (45:6), pp. 842-850.

Ariely, D., and Norton, M. I. 2007. "Psychology and Experimental Economics: A Gap in Abstraction," *Current Directions in Psychological Science* (16:6), pp. 336-339.

Bapna, R., Gupta, A., Rice, S., and Sundararajan, A. 2017. "Trust and the Strength of Ties in Online Social Networks: An Exploratory Field Experiment," *MIS Quarterly* (41:1), pp. 115-130.

Bapna, R., Ramaprasad, J., Shmueli, G., and Umyarov, A. 2016. "One-Way Mirrors in Online Dating: A Randomized Field Experiment," *Management Science* (62:11), pp. 3100-3122.

Bello, D., Leung, K., Radebaugh, L., Tung, R. L., and van Witteloostuijn, A. 2009. "From the Editors: Student Samples in International Business Research," *Journal of International Business Studies* (40:3), pp. 361-364.

Belmont Report. 1979. "Ethical Principles and Guidelines for the Protection of Human Subjects of Research," National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, U.S. Department of Health & Human Services.

Benbasat, I., and Zmud, R. W. 1999. "Empirical Research in Information Systems: The Practice or Relevance," *MIS Quarterly* (23:1), pp. 3-16.

Benlian, A., Titah, R., and Hess, T. 2012. "Differential Effects of Provider Recommendations and Consumer Reviews in E-Commerce Transactions: An Experimental Study," *Journal of Management Information Systems* (29:1), pp. 237-272.

Brewer, M. B., and Crano, W. D. 2000. "Research Design and Issues of Validity," in *Handbook of Research Methods in Social and Personality Psychology*. pp. 3-16.

Burtch, G., Hong, Y., Bapna, R., and Griskevicius, V. 2017. "Stimulating Online Reviews by Combining Financial Incentives and Social Norms," *Management Science* (64:5), pp. 2065-2082.

Campbell, D. T., and Stanley, J. C. 1966. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Company.

Cason, T. N., Kannan, K. N., and Siebert, R. 2011. "An Experimental Study of Information Revelation Policies in Sequential Auctions," *Management Science* (57:4), pp. 667-688.

Charness, G., Gneezy, U., and Kuhn, M. A. 2012. "Experimental Methods: Between-Subject and within-Subject Design," *Journal of Economic Behavior & Organization* (81:1), pp. 1-8.

Coolican, H. 2017. *Research Methods and Statistics in Psychology*. London: Psychology Press.

Cwiakowski, P., Giergiczny, M., and Krawczyk, M. 2016. "Pirates in the Lab: Using Incentivized Choice Experiments to Explore Preference for (Un) Authorized Content," *MIS Quarterly* (40:3), pp. 709-715.

Deb, A., Bhattacharya, S., Gu, J., Zhou, T., Feng, E., and Liu, M. 2018. "Under the Hood of Uber's Experimentation Platform." https://eng.uber.com/xp/.

Dennis, A. R., and Valacich, J. S. 2001. "Conducting Experimental Research in Information Systems," *Communications of the Association for Information Systems* (7:1), pp. 1-41.

Dipboye, R. L., and Flanagan, M. F. 1979. "Research Settings in Industrial and Organizational Psychology: Are Findings in the Field More Generalizable Than in the Laboratory?," *American Psychologist* (34:2), pp. 141-150.

Durgin, F. H., Baird, J. A., Greenburg, M., Russell, R., Shaughnessy, K., and Waymouth, S. 2009. "Who Is Being Deceived? The Experimental Demands of Wearing a Backpack," *Psychonomic Bulletin & Review* (16:5), pp. 964-969.

Ellis, T. J., and Levy, Y. 2009. "Towards a Guide for Novice Researchers on Research Methodology: Review and Proposed Methods," *Issues in Informing Science & Information Technology* (6), pp. 323-337.

Erlebacher, A. 1977. "Design and Analysis of Experiments Contrasting the within-and between-Subjects Manipulation of the Independent Variable," *Psychological Bulletin* (84:2), pp. 212-219.

Feenstra, R. 2015. "How Ice Cream Kills: Understanding Cause and Effect ". https://www.decisionskills.com/blog/how-ice-cream-kills-understanding-cause-and-effect.

Feitelson, D. G., Frachtenberg, E., and Beck, K. L. 2013. "Development and Deployment at Facebook," *IEEE Internet Computing* (17:4), pp. 8-17.

Fontenelle, G. A., Phillips, A. P., and Lane, D. M. 1985. "Generalizing across Stimuli as Well as Subjects: A Neglected Aspect of External Validity," *Journal of Applied Psychology* (70:1), pp. 101-107.

Fu, J.-R. 2011. "The Impact of Seller Expertise and a Refund Guarantee on Auction Outcome: Evidence from an Online Field Experiment of Camera Lens Market," *International Journal of Electronic Business Management* (9:3), pp. 187-195.

Greenwald, A. G. 1976. "Within-Subjects Designs: To Use or Not to Use?," *Psychological Bulletin* (83:2), pp. 314-320.

Gregg, D. G., and Walczak, S. 2008. "Dressing Your Online Auction Business for Success: An Experiment Comparing Two Ebay Businesses," *MIS Quarterly* (32:3), pp. 653-670.

Grimmelmann, J. 2015. "The Law and Ethics of Experiments on Social Media Users," *Colo. Tech. LJ* (13), pp. 219-271.

Gupta, A., Kannan, K., and Sanyal, P. 2018. "Economic Experiments in Information Systems," *Management Information Systems Quarterly* (42:2), pp. 595-606.

Harrison, G. W., and List, J. A. 2004. "Field Experiments," *Journal of Economic Literature* (42:4), pp. 1009-1055.

Hashim, M. J., Kannan, K. N., and Maximiano, S. 2017. "Information Feedback, Targeting, and Coordination: An Experimental Study," *Information Systems Research* (28:2), pp. 289-308.

Hauser, D. J., Ellsworth, P. C., and Gonzalez, R. 2018. "Are Manipulation Checks Necessary?," *Frontiers in Psychology* (9:998), pp. 1-10.

Hertwig, R., and Ortmann, A. 2008. "Deception in Experiments: Revisiting the Arguments in Its Defense," *Ethics & Behavior* (18:1), pp. 59-92.

Hui, K.-L., Teo, H. H., and Lee, S.-Y. T. 2007. "The Value of Privacy Assurance: An Exploratory Field Experiment," *MIS Quarterly* (31:1), pp. 19-33.

Kamis, A., Koufaris, M., and Stern, T. 2008. "Using an Attribute-Based Decision Support System for User-Customized Products Online: An Experimental Investigation," *MIS Quarterly* (32:1), pp. 159-177.

Karahanna, E., Benbasat, I., Bapna, R., and Rai, A. 2018. "Editor's Comments: Opportunities and Challenges for Different Types of Online Experiments," *MIS Quarterly* (42:4), pp. iii-x.

Keren, G., and Lewis, C. 1993. *A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Kohavi, R., Deng, A., Frasca, B., Walker, T., Xu, Y., and Pohlmann, N. 2013. "Online Controlled Experiments at Large Scale," *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: ACM, pp. 1168-1176.

Levine, C., Faden, R., Grady, C., Hammerschmidt, D., Eckenwiler, L., and Sugarman, J. 2004. "The Limitations of “Vulnerability” as a Protection for Human Research Participants," *The American Journal of Bioethics* (4:3), pp. 44-49.

Levy, Y., and Ellis, T. J. 2011. "A Guide for Novice Researchers on Experimental and Quasiexperimental Studies in Information Systems Research," *Interdisciplinary Journal of Information, Knowledge, and Management* (6), pp. 151-161.

Mangalaraj, G., Nerur, S., Mahapatra, R., and Price, K. H. 2014. "Distributed Cognition in Software Design: An Experimental Investigation of the Role of Design Patterns and Collaboration," *MIS Quarterly* (38:1), pp. 249-274.

Morris, S. B. 2008. "Estimating Effect Sizes from Pretest-Posttest-Control Group Designs," *Organizational Research Methods* (11:2), pp. 364-386.

Orne, M. T. 1962. "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications," *American Psychologist* (17:11), pp. 776-783.

Price, P. C., Jhangiani, R., and Chiang, I.-C. A. 2017. *Research Methods in Psychology*. BCCampus.

Rice, S. C. 2012. "Reputation and Uncertainty in Online Markets: An Experimental Study," *Information Systems Research* (23:2), pp. 436-452.

Rosnow, R. L., and Suls, J. M. 1970. "Reactive Effects of Pretesting in Attitude Research," *Journal of Personality and Social Psychology* (15:4), pp. 338-343.

Rubin, D. B. 2008. "For Objective Causal Inference, Design Trumps Analysis," *The Annals of Applied Statistics* (2:3), pp. 808-840.

Rutherford, A. 2001. *Introducing Anova and Ancova: A Glm Approach*. Thousand Oaks, CA: SAGE Publications Ltd.

Scheffe, H. 1999. *The Analysis of Variance*. Hoboken, NJ: Wiley-Interscience.

Shadish, W. R., Cook, T. D., and Campbell, D. T. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Suresh, K. 2011. "An Overview of Randomization Techniques: An Unbiased Assessment of Outcome in Clinical Research," *Journal of Human Reproductive Sciences* (4:1), pp. 8-11.

Sutanto, J., Palme, E., Tan, C.-H., and Phang, C. W. 2013. "Addressing the Personalization-Privacy Paradox: An Empirical Assessment from a Field Experiment on Smartphone Users," *MIS Quarterly* (37:4), pp. 1141-1164.

Tams, S., Thatcher, J. B., and Grover, V. 2018. "Concentration, Competence, Confidence, and Capture: An Experimental Study of Age, Interruption-Based Technostress, and Task Performance," *Journal of the Association for Information Systems* (19:9), pp. 857-908.

Tang, D., Agarwal, A., O'Brien, D., and Meyer, M. 2010. "Overlapping Experiment Infrastructure: More, Better, Faster Experimentation," *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: ACM, pp. 17-26.

Thatcher, J. B., Wright, R. T., Sun, H., Zagenczyk, T. J., and Klein, R. 2018. "Mindfulness in Information Technology Use: Definitions, Distinctions, and a New Measure," *MIS Quarterly* (42:3), pp. 831-847.

Tsai, J. Y., Egelman, S., Cranor, L., and Acquisti, A. 2011. "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study," *Information Systems Research* (22:2), pp. 254-268.

Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M., and Marett, K. 2014. "Research Note—Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance," *Information Systems Research* (25:2), pp. 385-400.

Yan, E. G., and Munir, K. M. 2004. "Regulatory and Ethical Principles in Research Involving Children and Individuals with Developmental Disabilities," *Ethics & Behavior* (14:1), pp. 31-49.