BRG REVIEW

SPRING 2013 Volume 3, Issue 1

BRG BERKELEY RESEARCH GROUP

BRG REVIEW

Volume 3 Issue 1

BRG REVIEW

Subscription. For subscription questions, problems, or address changes:

The BRG Review
Berkeley Research Group, LLC
2049 Century Park East, Suite 2525
Los Angeles, CA 90067
510.550.4380
info@brg-expert.com

Disclaimer. The opinions expressed in the BRG Review are those of the individual authors and do not represent the opinions of BRG or its other employees and affiliates. The information provided in the BRG Review is not intended to and does not render legal, accounting, tax, or other professional advice or services, and no client relationship is established with BRG by making any information available in this publication, or from you transmitting an email or other message to us. None of the information contained herein should be used as a substitute for consultation with competent advisors.

ISSN 2326-0653 (print)
ISSN 2168-9814 (online)

BRG REVIEW

Table of Contents

From the Desk of the Editor

Welcome to the fourth issue of the *Berkeley Research Group Review*, the official publication of Berkeley Research Group, LLC. BRG was founded in 2010 by a group of distinguished academics and private sector professionals in the fields of economics, finance, healthcare, and data analytics.

BRG's expertise includes antitrust and competition policy, bankruptcy and insolvency, class action litigation, damages analysis, data analytics, finance and valuation, forensic accounting and investigations, healthcare, intellectual property, public policy, and securities.

In this issue, we are pleased to present two case studies from BRG's Labor and Employment practice. I have asked Professor David Lewin (UCLA), the head of our labor practice, to introduce the papers.

We hope to use the *Review* to provide our audience with a "good read" and improve our connections with clients, recruits, peers, and colleagues. We expect that the *Review* will stimulate discussion and debate around key issues we face today. With this in mind, we welcome any comments or feedback you have about the subjects we raise in the *Review.*

Kindest regards,

Introduction to Observational Studies Articles for the BRG Review

David Lewin

Expert work in wage and hour litigation often requires the use of primary research methods—in particular, surveys, interviews, and observational studies. These methods are aimed at determining the work—job tasks—performed by employees and/or the amount of time employees spend performing their work. Such primary research has been mainly spurred by litigation in which it is alleged that front-line managers and supervisors largely perform employee work and that work should therefore be reclassified as non-exempt under federal and state wage and hour laws.[1] It has also been spurred, however, by litigation involving non-managerial employees, such as claims adjusters and sales representatives, who have traditionally been regarded and classified by employers as exempt from these same federal and state wage and hour laws.[2]

While surveys and interviews can help in determining employee job tasks and the amount of time employees spend performing such tasks, both methods are limited by the fact that they basically elicit opinions from survey and interview subjects. Such opinions, whether obtained from current or former employees, are open to challenge in terms of validity and reliability. For example, current employees who are parties to wage and hour litigation may over-report the time they spend performing what is in essence non-exempt-type work. In another example, former employees who are parties to wage and hour litigation are basically asked to recall from memory the job tasks they performed and amount of time they spent performing those tasks, which presumes no memory decay.

1 This is commonly referred to as "managerial misclassification" litigation. For an analysis, see D.I. Levine and D. Lewin, "The New Managerial Misclassification Challenge to Old Wage and Hour Law; Or, What is Managerial Work?" In D. Lewin (Ed.), *Contemporary Issues in Employment Relations* (2006). Champaign, IL: Labor and Employment Relations Association, pp. 189–222.

2 For one such example, see *Nettles, Czarnecki et al. vs. Allstate Insurance Company* (2009).

For these and related reasons, observational studies are preferable to survey and interview methods for determining the job tasks performed by employees and the amount of time they spend performing those tasks.[3] Observational studies do not require reliance on employee opinions in these respects; rather, they rely on first-hand observations of employees actually performing their work and the amount of time taken to perform that work. The validity and reliability of such observations is likely to be substantially greater than the validity and reliability of employee opinions obtained through surveys and interviews.

With these considerations in mind, this issue of the *BRG Review* features two articles—the first by Stefan Boedeker and the second by Dawn Eash—that report on recent observational studies Berkeley Research Group conducted in the context of wage and hour litigation.[4] Boedeker's article provides a detailed account of an observational study conducted at an amusement park that was sued by supervisors and assistant managers who claimed that they were erroneously classified as exempt employees because they performed essentially the same work as non-exempt employees and should therefore be eligible for overtime pay. Particularly notable, Boedeker frames this case study with careful analysis of key factors to consider in designing an observational study (e.g., subjects' knowledge of their being observed) and in obtaining and maintaining high-quality observational data.

Eash's article provides a comprehensive account of an observational study conducted for a company whose sales representatives were deemed by a court

3 An observational study is not the same as—and not a synonym for—a time and motion study. In an observational study, the main objectives typically are to determine the job tasks that employees perform and the amount of time they spent performing those tasks. In a time and motion study, these same objectives are in play, but additional objectives are to set standard times for the performance of job tasks and to specify prevailing rates of pay for meeting, exceeding, and falling short of such standard times. These latter objectives are not operative in an observational study conducted in the context of litigation.

4 Another example of an observational study conducted by BRG, which California courts subsequently relied on in rendering decisions in a series of managerial misclassification cases involving a retail company, is *Keller, et al. vs. Tuesday Morning, Inc.* (2007).

to be covered by (i.e., non-exempt under) federal and state labor laws and were hence eligible for overtime pay. The key issue addressed by this study is the amount of time that sales representatives spend performing their work—something that the pharmaceutical company had never recorded or measured. This study, which involved repeat observations of a stratified random sample of sales representatives, is perhaps the largest ever conducted for litigation or research purposes.

Taken together, these two articles attest to the capability offered by BRG in conducting observational studies in the context of wage and hour litigation.

Stefan Boedeker
Berkeley Research Group

Stefan Boedeker specializes in the application of economic, statistical, and financial models to areas such as solutions to business issues, economic impact studies, and complex litigation cases.

With over 20 years of experience, Mr. Boedeker has provided economic, financial, and statistical consulting and expert services to clients across a wide range of industries, including healthcare, high technology, entertainment, manufacturing, retail, real estate, insurance, and financial services, and federal, state, and local governments. Mr. Boedeker has issued expert reports and given deposition and trial testimony in state and federal courts.

Mr. Boedeker has extensive experience applying economic and statistical theories to class action-related matters, including class certification issues, liability assessment, and calculation of economic damages.

Mr. Boedeker also has extensive experience applying economic and statistical theories and methodologies to employment-related matters such as discrimination, wrongful termination, and wage and hour cases. His work in such cases to date has included designing and conducting surveys, time and motion studies, and observational studies; statistically analyzing the results of such surveys and studies; applying statistical sampling methodologies to extrapolate results from a subset to a universe of individuals; developing statistical models and tests to answer liability questions; applying economic theory to develop damages scenarios; and analyzing large employment-related databases.

Contact Information:
Email: sboedeker@brg-expert.com
Phone: 310.499.4924

Observational Studies in Wage and Hour Litigation

Abstract

In many wage and hour class actions, questions of liability can only be answered by analyzing how employees spend their workdays. Typically, employers and employees do not keep detailed logs about daily activities that could be used to answer those questions. Observational studies are important tools to collect data that are not otherwise available but that experts need to render opinions. In this paper, we discuss steps to take in the design, implementation, and statistical analysis phases of an observational study to ensure that the study generates high-quality data that the expert and trier of facts can rely on in the course of the litigation. We conclude the paper with a summary of an actual case study.

I. *Introduction*

A wave of wage and hour litigation cases emerged in the 1990s; many were filed in California, but they quickly spread to the rest of the nation. These cases dealt mainly with allegations of employee misclassification, under- or non-payment of overtime, off-the-clock work, meal and rest break violations, under- or non-payment of donning and doffing activities, unlawful vacation policies, etc. In many wage and hour cases, two questions are crucial in determining liability and damages: (1) How do employees spend their time qualitatively? (i.e., what kind of activities do employees engage in during their workday?); and (2) How do employees allocate their time quantitatively across different tasks during their workday?

Early in the history of wage and hour litigation cases, consultants and experts "borrowed" two data-collection approaches from the field of industrial engineering to find reliable and defensible answers to these questions: (1) time

and motion studies and (2) observational studies. In time and motion studies, certain tasks and work processes are isolated, and the duration of the tasks and processes in the course of employees' work shifts, days, or weeks are repeatedly measured until enough observations have been collected to obtain an average standard time. By contrast, observational studies focus on entire shifts, workdays, or work weeks of selected employees to obtain qualitative information about the kinds of tasks they engage in and quantitative measurements of the duration of those tasks.

Depending on the requirements in a particular case, either data-collection method may yield useful results. For example, while a time and motion study would be appropriate in a donning and doffing case, an observational study would be appropriate in a managerial misclassification case. This article focuses on the importance of data quality in observational studies and how to achieve high-quality data by laying the proper planning and design process foundation and maintaining quality in the data management and analysis phases of a study. A brief case study provides an illustrative example.

II. *General Considerations*

Wage and hour litigation began mainly in the fast food industry and more broadly in the retail sector (e.g., big box retailers, department stores), but quickly spread to almost every industry and sector. There are many ways to design an observational study. Each wage and hour litigation case may be guided by a slightly different objective, as dictated by the allegations in the complaint. For example, a case alleging manager misclassification requires a different approach than a case alleging meal and rest break violations. Further, the industry involved in a particular case has a significant impact on the research design of the observational study. For example, a case involving employees in gas station convenience stores requires a different design than a case involving sales associates in specialty accessory stores located in malls.

The type and format of data that need to be collected are key drivers in the design phase and the actual implementation of every observational study. In addition to these external factors, internal factors such as time and budget constraints have to be considered. Given these considerations, a brief discussion of six important features of observational studies is presented below. These features should be carefully taken into account when contemplating using an observational study as a data-collection tool in a wage and hour case.

Study Setting

A study should be conducted in the "real life" environment where employees actually perform their work. This simply means observing subjects in the work areas where they spend their shifts, days, or weeks. Ideally, observations can be made such that the observers do not influence the work performed by subjects, such as by slowing it down or otherwise negatively impacting related work processes.

Knowledge of Being Observed

When subjects do not know they are being observed, the study can be classified as "disguised." In such a study, subjects are highly likely to act as naturally as they would if they were not being observed. Thus, the data collected from this type of study tend to reflect employees' "true" work and "true" allocation of time among different tasks. Sometimes, however, disguised studies raise ethical concerns about collecting and recording employees' work-related behavior without their consent.

By contrast, a study is referred to as "non-disguised" when subjects know they are being observed. Using the non-disguised method alleviates ethical concerns, but non-disguised studies have to be designed such that the observations are not biased by the Hawthorne Effect. This refers to the phenomenon whereby subjects who know they are being observed alter their behavior to a point where the observations no longer adequately reflect their true work and true allocation of time. Therefore, it is important in designing a

non-disguised observational study to include repeat observations of the same subjects and of subjects in similarly situated jobs.

Data-collection Methodology

In general, observational methods range from entirely "human" to entirely "mechanical." Traditionally, time and motion studies and observational studies were performed by industrial engineers equipped with a clipboard and stop watch, which they used while following employees at the workplace and observing them performing their work tasks and processes. Observers frequently used this traditional method of collecting data in early wage and hour litigation-related studies. Since then, and with the rapid increase of relevant technology, observational tools have become more sophisticated. Today, observers often use iPads and/or digital cameras to collect data. Proprietary software is available to digitally analyze video footage, measure elapsed times between work-related tasks and events, and tabulate the resulting data in a spreadsheet format. If the location specifics permit, observations are recorded by preinstalled video systems, and human observers are eliminated from the study. This method reduces the Hawthorne Effect—preinstalled systems are less intrusive than human observers. Video systems also offer a more precise method of data collection.

Context of the Observed Activity

A study subject may be observed stocking shelves, but without recording the context of this observation, it is difficult if not impossible to assess the true nature of the activity. From a potential liability and damages perspective, for example, there is a significant difference between a) an observed assistant manager who stocks shelves because no one else is available and b) an assistant manager who stocks shelves in order to show and explain to a group of new employees the process of correct product placement. Studies that fail to capture context either by recording it in addition to the tasks themselves

and the duration of the tasks, or by capturing it on camera, yield much less reliable, defensible results.

Degree of Structure Imposed by Data-collection Method

At one extreme, completely structured observations are made when the data being collected have been organized into clearly defined categories or groups so that the observer can record the data by simply marking off or checking a category on an observation form and recording the respective elapsed times. At the other extreme, no structure is imposed, and the observer's only task is to capture the entire work shift or day either on camera or through some other medium.

Adding "Staged" Observations to "Real Life" Observations

As discussed above, an observational study should capture the real-life situation in which the study subjects perform their work. However, it is entirely possible that the subjects in a non-disguised setting alter their behavior to a degree that the measurements no longer truly represent their work tasks and allocation of time to different tasks. In these instances, a "staged" setting can be utilized, whereby the observer creates the specific situation being studied. For example, a subject who takes 30 minutes to restock a shelf may have performed the task much more slowly than usual because he knows that he is being observed. In a staged setting, an individual who is not part of the initial study group (typically the class) can be observed performing the same task after hours. The results from a sufficiently large number of staged observations can function as a benchmark or surrogate for observations of real-life situations.

III. ***Achieving Data Quality in Observational Studies***

The following sections contain a brief discussion of the importance of achieving data quality in observational studies; and a detailed discussion of how to achieve data quality in observational studies through careful,

thorough planning in the design phase and by applying scientific rigor and sound methodology in the implementation phase.

A. *The Importance of Achieving Data Quality*

It is important to point out that the scope and duration of observational studies are quite limited in two respects—the number of subjects observed and the time period covered. A class period typically covers multiple years and often includes as plaintiffs thousands of former and current employees, whereas an observational study is typically performed in a few days or weeks and covers a relatively small number of individuals. By virtue of this limitation, every data point collected in an observational study carries the weight of thousands of events that were unobserved during the class period.

Observational studies also must be designed to adequately represent the real-life situations that potential class members have experienced while working in an environment defined by the alleged wrongdoing of the employer. Such situations are often quite messy and hard to capture in a precise way.

The results of an observational study ultimately have to be relied upon by the trier of fact or by the parties in the dispute during settlement negotiations. Any potential doubt about the quality of the study makes its usefulness as a decision support tool questionable and, in the worst case, unusable. Therefore, the study must be designed and implemented such that data bias is ruled out or at least minimized. Further, the data reporting process and documentation must be transparent, and an audit trail has to be established to enable others to replicate results of the study.

B. ***Laying the Foundation in the Planning and Design Phase***

Here we provide an in-depth discussion of challenges to achieve data quality in the planning and design phase of an observational study. Points of particular importance are:

- Define the population of interest
- Identify auxiliary data sources
- Identify the selection process of study subjects
- Define measurement methodology
- Incorporate controls to avoid sampling bias, observer bias, and measurement error
- Select sites and actual timing of the study

The first challenge lies in the difficulty of precisely defining the population of interest (i.e., the class), because it typically consists of current and former employees. Any observational study is necessarily limited to the accessible population, which is the group of current employees in the positions for which alleged wrongdoing by the employer may have caused economic damage. Before any decisions about location, duration, and sample size can be made, a thorough analysis has to be performed of the work environment experienced by employees during the alleged class period.

The identification of additional historical business, operational, and human resource data may be necessary to show that the intended scope of a study performed in the present in terms of location, duration, and employee coverage can yield results representative of the entire class period. Without those auxiliary data sources, the study will be open to criticism about how a study performed in the present can be used to draw valid and reliable inferences about employees' work-related experiences at the beginning of the class period, which in many cases may be four or more years earlier.

Next, the question arises as to how many employees to observe and how to select the study subjects. Typically, studying an entire population is time and

cost prohibitive; therefore, sampling methodology has to be applied. Ideally, information about current business locations or sites (e.g., stores, warehouses, routes) can be used to select a statistically valid random sample of locations/sites and the number of study subjects per site. However, even in the ideal case of a fully statistically valid random sample, the population of interest may change between the time of planning the study and actually conducting the study (employees may quit or be discharged, get promoted, change locations, change work status, etc.). The results from an observational study often cannot be extrapolated to the underlying population of interest with a known degree of statistical confidence and precision.

If a statistically random sample design is not possible, or if a hybrid is utilized (i.e., a two-stage process in which locations/sites are randomly selected in the first stage, and study subjects are selected either by random selection or observer choice in the second stage), it is important to ensure that the selection process yields study subjects who are representative of the population of interest. Representativeness can often be achieved through the application of a quota system that, if applied correctly, ensures that the relative frequency of a stratum (of jobs and/or employees) within the universe is approximately the same as the relative frequency of a stratum within the sample. Well-designed, non-random quota samples with a sufficiently large sample size can yield useful and reliable results even though statistical confidence and precision cannot be fully measured.

Sample size often can be addressed through a pilot or pre-study. In general, the sample size in statistically random samples depends on the desired precision and confidence levels and on the underlying variation in the variable of interest. For example, if a case alleges under- or non-payment of overtime hours worked, then an environment in which the observed overtime worked by employees ranges from 0 to 5 hours per week requires a smaller sample size than an environment in which the observed overtime worked by employees ranges from 0 to 20 hours per week. Apart from these considerations, budget and time constraints frequently prohibit the selection of a large sample

of employees to be observed. In such instances, it is especially important to ensure that each data point collected satisfies the highest quality standard.

Additional considerations in selecting location/site and employee/subject samples for observational study include multi-shift operations, seasonality, weekend and holiday work, and employee turnover. A large sample is good, more is better, and much more is still better. Therefore, spreading available resources (time, observers, and budget) across as many different study subjects as possible is preferable to collecting more information on fewer study subjects. Nonetheless, careful consideration has to be given to capture the broad spectrum of tasks a study subject may perform or be involved in. For example, a store manager may have recurring daily, weekly, and monthly duties that render an observational study of what she does on a Wednesday incomplete and thus unreliable. In this instance, observation days should be selected to take account of how the store manager's weekly and monthly recurring tasks impact her allocation of time. Similarly, observing the store manager only on a day when month-end reports are due will likely give an incorrect assessment of what she typically does.

At this point in the planning and design phase, data measurement scales must be determined. Data measurements recorded during the study affect data quality, which ultimately affects the choice of statistical analysis methodology and the strength and reliability of conclusions that can be drawn from the results. It is preferable to plan for more detailed data and granularity, because these data can be summarized or collapsed into categories; however, the converse is not true. If data have been collected and recorded, but it turns out that the level of detail was insufficient, it is too late to go back to the drawing board and redesign the study.

In sampling-based data-collection methods, an inherent sampling error stems from the fact that not all members of the universe have been selected for observation. A different sample may have yielded different results, which would cast doubt on the results, especially when the study deviates from the ideal random-sampling approach. Therefore—and if possible—a pilot or pre-study

should be conducted in order to determine an adequate sample size. If the selection of study subjects is based on non-random selection criteria, a quota system has to be established to ensure that the selected study subjects can be considered representative of the underlying universe.

The presence of systematic or non-sampling bias is a potential related problem. Sample selection bias can skew results when no quota systems exist or when study subjects are selected based on convenience and availability. Non-random samples that are not representative of the universe will yield unreliable results.

In many observational studies, data quality suffers from untruthful or exaggerated behavior. It is important to build in measures to counter the impact of such behavior, either by staging certain activities or by having repeated observations of the same subjects or additional observations of alternate subjects. Two other issues negatively impact data quality: potential observer bias and measurement error. Observer bias can occur when the observer has a desired outcome in mind and consciously or unconsciously takes or records his measurements to achieve that outcome. Improper use of equipment or faulty equipment can cause measurement error. To avoid these biases, observers should be trained in observational methods, instructed to avoid preconceived notions about potentially desired outcomes, and shown how to properly use observational recording equipment.

Another key aspect of an observational study is timing. If observational data have to be produced quickly, then all observations may have to be collected in one contiguous visit per site. If extraneous factors significantly impact work tasks and processes, multiple visits to the same site may be necessary in order to capture the variation caused by such factors.

The feasibility of an observational study has to be performed before beginning the study. The following questions can help determine such feasibility:

- Does the design meet the objectives of the legal team and the end client?
- Can the design meet timeline and budget expectations?
- Is the targeted sample size feasible when considering:
  - Specific inclusion and exclusion criteria and quotas?
  - Study subject availability?
  - Sampling methodology, including need for potential stratification?
  - Adequate precautions and replacements for unexpected drop out of study subjects?
- Are enough trained and qualified observers available?
- Are the appropriate operational resources (i.e., project management, data management, site management, technology, etc.) available to successfully implement the study design?

C. *Maintaining Data Quality in the Data Management and Analysis Phase*

As previously discussed, the main purpose of an observational study is to serve as a data-collection tool. The study design phase should include a data management plan. Depending on where the proposed study falls on the spectrum between purely human- (clipboard and stopwatch) and purely mechanical- (preinstalled video systems) type observational studies, the design phase should ensure that collected and recorded observations are properly transferred into a format that can be utilized for data analysis. For example, in a study in which data are recorded based on human observation, hand-taken time measurements, comments, and notes, a protocol must be established for transcribing the raw data into an electronic data format (i.e.,

spreadsheet, database, etc.). Data formatting and standardization must be documented so that a seamless audit trail from originally collected data to generation of the data for statistical analysis can be produced.

In the case where a technology-supported human observation technique is utilized (e.g., observers use iPads), a protocol must be established to transfer the data from the originally used equipment to the system that will ultimately analyze the data. Ideally, algorithms for data transfer and data formatting can be saved in the form of batch files of computer commands that can be repeatedly executed to replicate the data used for analysis. In instances where video equipment is used to record the work performed by employees, a protocol has to be established to transfer the footage into a system that can generate qualitative and quantitative data.

Regardless of the data-collection methodology used, it is important to begin the data review process concurrent with the data-collection process. Ideally, data from an observed work shift or day are extracted from the data-collection tool at the end of each observation increment and transferred to a centralized data storage and data management system. The timely summary and transfer of such data becomes even more important when multiple observers in different locations are involved. Early and frequent review of data will contribute to a better understanding of the data-generation process, help identify trends, and assist in estimating how many more observations may be necessary to obtain reliable results. The use of a centralized data management system also enables the consistent application of targeted data-cleansing procedures, plausibility checks, and quality-control queries. This aspect of concurrent quality control plays an important role when potential problem areas in the data-collection process may have been overlooked or misinterpreted in the design phase. The central quality-control function may also identify issues and ways of dealing with them through specific training before they can compromise the integrity and validity of the study.

Finally, a detailed plan for statistically analyzing the data should be developed during the design phase. Appropriate statistical measures to capture the central tendency and dispersion of the observed data distributions have to be calculated at regular intervals as data are collected. The statistical analysis plan should also contain provisions regarding how to treat potentially missing or incomplete data and implausible data or outliers. Missing or incomplete data can arise quite frequently in observational studies when preselected study subjects are not available, schedules change, and logistical needs require location changes. In addition, observer error and mistakes or equipment malfunction and/or failure may render certain observations unusable. The treatment of implausible observations or outliers is best dealt with at the outset of a study to mitigate or avoid subsequent criticism. It is not unusual to propose the use of a trimmed distribution rather than using every observation to report final study results. Depending on the shape of the distribution of the observed data (e.g., bell shaped or right-tail skewed), a trimmed distribution excludes specific percentages of observation in both tails (e.g., when calculating summary statistics in a bell-shaped data distribution, the trimmed distribution approach excludes the lowest and highest 5 percent of data points).

IV. *Case Study*

An operator of large amusement parks was sued by supervisors and assistant managers in its Foods Administration division over misclassification of their status as exempt employees. The complaint alleged that these employees essentially performed the same job duties and tasks as the hourly food workers without getting paid for overtime work. BRG was retained to conduct an empirical study of daily activities of supervisors and assistant managers in the division at the largest of the parks. The scope of BRG's retention was an observational study, because the objective was to determine how supervisors and assistant managers allocate their time across different tasks during a regular work week.

The study was designed to observe, record, and analyze the daily activities of Foods Administration supervisors and assistant managers at the selected parks. Observations were obtained using two approaches recording the entire shift of an employee: 1) via handheld camcorders operated by trained observers, and 2) via fixed installed surveillance cameras. By recording the complete daily shift for an entire work week, any attempt by subjects to alter their behavior knowing that they are being recorded on camera is minimized in comparison to just recording short-time increments or certain tasks. Even though the subjects are aware of the cameras, they eventually must engage in activities necessary to do their work.

Observational studies based on video footage over extended time periods are far superior to surveys and questionnaires as data-collection tools because they can be objectively timed, and the actual activities are captured as they occur. Assessments and definitions of tasks can be modified and improved as video footage is reviewed and coded; if necessary, footage can be reviewed multiple times. In contrast, surveys and interviews rely on respondents' memories to evaluate and summarize events that have occurred in the past—sometimes many years in the past. Also with surveys and interviews, respondents are typically given long task lists and then asked to estimate how much time they spent performing these tasks in a typical work week. Additionally, these respondents are often asked to make certain mathematical computations, such as aggregating the amount of time spent performing tasks during a particular time period and then averaging the time or applying percentages to hours worked.

Observational studies have a disadvantage: they often only cover a limited amount of time, such as a work week. Questions may arise as to whether the selected week adequately reflects work activities in order to draw conclusions about the allocation of time across tasks for the entire tenure of an individual employee, and across all employees in a class list during the entire class period. However, studies can overcome this potential disadvantage by analyzing

auxiliary data that can shed light on the representative nature of the selected time window that was observed.

For this observational study, such auxiliary data existed: a) work-scheduling data for the class members, b) timekeeping and payroll records of employees' work hours, c) statistics of daily park visitors, and d) total visitor spending for food. These data were used to specify a statistical model to test whether the activities of food workers in an amusement park correlate with how busy the park is on a given day and how many individuals are scheduled to work. The choice of study time period was determined by external factors, including the points in time when BRG was retained and when the study results had to be rendered. If the statistical model utilizing these auxiliary data showed significant results, then the selected observation period would be an adequate representation of the overall time period of interest. In addition, more qualitative variables—such as job descriptions, job titles, job hierarchy, park layout, and locations of food establishments—were analyzed to assess whether these factors changed during the class period, with potential consequent changes in workflow and/or allocation of time across different tasks. Analysis of these qualitative variables indicated no changes within the class period.

The study was conducted between Monday, August 1, 2011, and Monday, September 5, 2011 (Labor Day weekend). To determine whether the study was representative of the overall class period, statistical models were specified to explain the relationships among food services revenue, park attendance, the number of food workers scheduled to work, and the numbers of supervisors and assistant managers scheduled to work. The first regression analysis considered the entire time period for which the data were available (i.e., the 42nd week of 2004 through the 39th week of 2011). Weeks 31 through 35 of 2011 cover the period when the study was performed (August 1 to September 5, 2011). Multivariate logarithmic regression models were specified with the supervisor- and assistant manager–staffing level as the dependent variable and park attendance, food establishment revenue, and staffing levels of food workers as independent variables. Statistically significant correlations

(i.e., regression coefficients) were found between revenue at food establishments and park attendance and staffing levels at food establishments.

Moreover, the results indicated that the month of August 2011 was not statistically significantly different from the remainder of the observations from late October 2004 to the end of July 2011. The R-squared, which measures the percentage of variation in the dependent variable (i.e., the staffing level of supervisors and assistant managers) that is explained by the model, was in excess of 93 percent. Other statistical tests indicated that the hypothesis that the regression coefficients were not significantly different from each other could not be rejected. In addition to the overall month-to-month regression models, a second set of regression analyses ware performed to test whether there were significant differences among individual years of the class period. These regressions indicated that revenue levels were a statistically significant predictor of staffing levels at food establishments and, moreover, that the year-to-year functional relationships were not statistically significantly different. In sum, the results of qualitative and quantitative auxiliary data analysis clearly indicated that the time period that was available to conduct the observational study was not statistically significantly different from other time periods during the class period.

In the next stage of the design and planning phase, a detailed site visit was conducted, during which test measurements were taken to determine where to place observers, identify the most appropriate observation style, assess scheduling details, and ultimately select the study subjects. Several distinctively different food establishments exist in the amusement park: company-owned sit-down restaurants, chain restaurants operating on the park premises, walk-up food stations with seating provided, walk-up food stations with no seating provided, and mobile carts. Separate study plans were developed for the indoor restaurants and outdoor walk-up stations. The observations were collected through observers operating handheld camcorders and through pre installed surveillance cameras. The existing surveillance camera system was enhanced by installing additional cameras in some indoor restaurants.

The observer-based part of the study was designed to use handheld camcorders to record assistant managers' and supervisors' daily activities in predefined outdoor "areas" of the park. The park-owned and operated restaurants were also covered by observers with handheld camcorders, because the site inspection revealed that these locations were too spread out and too large to efficiently use surveillance cameras.

A group of observers with camcorders was sent to the park for two one-week time increments, one of which covered Labor Day weekend. The week of Labor Day weekend also included some once-a-month, month-end administrative tasks that altered the allocation of time between tasks. The company's Human Resources department provided schedules that were used to randomly select assistant managers and supervisors such that each of the distinct areas was covered. The selected assistant managers and supervisors were each scheduled for five days of observation. The observational study covered a full seven-day week, which led to observation of replacements for the scheduled assistant managers and supervisors. (To illustrate, if Supervisor A was scheduled for an area Thursday through Monday, we also observed the supervisor who was scheduled for that area on Tuesday and Wednesday, which may have been the same individual replacing Supervisor A on both days or two different individuals).

Before the first day of observations, observers participated in a four-hour training session. This session covered the proper way to use the camcorders (e.g., keeping them turned on at all times to ensure an accurate representation of the amount of time spent on all tasks) and other instructions necessary to complete the study. Observers were given only very basic information about the study to ensure that no information was passed on to the subjects. A second training session was performed as an on-site walk-through so that observers had the opportunity to see the surroundings of their designated area prior to beginning their observations.

The Human Resources department provided the study subjects' schedules. Observers were instructed to arrive at the worksite at least 30 minutes before

the scheduled start time to ensure that potential issues could be dealt with by the time the actual shift started and the entire work day could be observed without problems. In addition, observers were given specific instructions regarding how they were to behave at the park and interact with their subjects. In particular, to ensure that subjects went about their normal daily routine as much as possible, observers were told not to engage in conversation with the subjects and not to interfere with their work in any way, with the exception being conversations regarding logistics (e.g., schedule changes and restroom breaks). Observers were also instructed not to share their schedule with the subjects, so that subjects were only informed of future observations the night before the observations took place. For example, an observer observing a subject on Monday, Tuesday, and Wednesday would inform their assigned subject of Tuesday's observation on Monday night and of Wednesday's observation on Tuesday night. This ensured that subjects could not artificially overload their schedules with tasks that may not have been representative of their typical activities. A supervising observer extracted all video footage from the camcorders at the end of each observation day and uploaded the footage on site into a central data storage hard drive. That hard drive was then taken to BRG's office and immediately uploaded into a secure case folder to perform initial steps of quality control, data standardization, and data formatting.

The employees in the chain restaurants were observed using surveillance cameras. In addition to existing surveillance systems, fixed surveillance cameras and a digital video recorder (DVR) were installed at each location. The cameras ran continuously during the study period. A supervising observer extracted all video footage from the hard drive of the surveillance system at the end of each observation day and uploaded the footage on site into a central data storage hard drive. As with the hard drive storing camcorder video footage, that hard drive was taken to BRG's office and uploaded into a secure case folder for quality control, data standardization, and data formatting.

Once the footage was in house at BRG, a group of video coders was tasked to review over 600 gigabytes of camcorder footage covering almost 500 hours

of observations, as well as more than 800 hours and thousands of gigabytes of surveillance video footage. Once the surveillance videos were limited to the relevant hours (based on the operating hours for each restaurant), 862 gigabytes of uncompressed video footage remained. The coders were charged with assigning task codes to employees' actions. A list of over 30 task codes had been defined and agreed upon by plaintiffs' and defense counsel. The video coding process took 19 days to complete and yielded slightly over 40,000 entries for the camcorder observations and almost 32,000 entries for the surveillance camera observations.

Before beginning the video coding, coders participated in a one-day training session to familiarize them with the videos, video coding software, and task list. The coders reviewed actual footage and were instructed on how to code specific tasks and sequences of tasks as the subject in the video switched from task to task and frame to frame. During the training sessions, instruction of how to code specific tasks was given to each coder for the same task. To achieve consistency, the process was repeated until each coder coded an observed task the same way. Once consistency had been achieved, the actual coding began. The coders had no background knowledge about the study, so as not to influence their thought process while coding. In addition, two quality control consultants worked full time with the coders in order to instantly review the work product.

Coders were instructed to pay close attention to video to ensure accurate coding of observed activities. In cases when the activity was not immediately apparent, coders were instructed to use context clues to help determine how to classify tasks, review the footage again, or both. For camcorder observations, coders could use audio context clues to help determine the nature of various tasks. For surveillance camera observations, coders were told to use a subject's surroundings and body language to aid in classifying tasks.

As some initial review of the video footage revealed, supervisors spent a significant amount of time during their daily shift walking the premises/areas that were under their supervision. Oftentimes the walking was interrupted

by brief activities lasting just a few seconds (e.g., picking up menus and placing them at the hostess stand, taking a check from a table and placing it at the cash register). If the activity was concluded before the start of the tenth second, the activity was not coded separately but instead summarized under the walking category. However (and for example), if a supervisor picked up a check and instead of just walking it over to the cash register actually rang it up at the register and completed the transaction, this entire activity was coded as cashiering.

Once coding was completed, statistical summaries were computed on an individual shift level, an aggregate level for each selected individual, and an aggregate level for each department. Summaries contained total times spent on all the activities the study subjects engaged in during the entire time they were observed. Summaries did not contain assessments of the status of an activity as "exempt" or "non-exempt."

It is notable that the task "Traveling/walking at/to site" was the individual activity with the most time spent for each individual study subject and across all study subjects. In fact, the study subjects spent 28.3 percent of the entire recorded time on this activity. Typically, this activity does not show up in job descriptions and, therefore, is often a subject of disagreement between plaintiff's and defense counsel about how to classify it. In my experience of many observational studies conducted in the food, retail, and manufacturing sectors, this task always takes up a large amount of a supervisor or manager's time. The question that typically comes up is how to classify the time spent on this activity. In this case, defense and plaintiff's counsel asked, "Could the study be used to correctly classify the task of 'Traveling/walking at/to site'?"

In general, some jobs for food workers in the restaurant business include traveling or walking at the worksite. Hostesses, waitresses, and busing personnel typically spend a lot of time walking at their respective worksites. However, the walking and traveling at the worksite as done by a hostess, waitress, or busing person is fundamentally different from that of a supervisor or assistant manager. The hostess, waitress, and busing person typically travel between a

very limited number of locations, engaging in a few activities at the start and end points of the walking activity. For example, a hostess always greets the guests at the hostess stand, walks them to their tables, and then walks back to the hostess stand. A waitress takes orders at the table and then walks to the kitchen or bar with the order, checks on her assigned tables, and rings up checks. A busing person collects the dirty dishes and then walks between the tables and the dishwashing/cleaning area.

In contrast, the walking activity of a supervisor or assistant manager covers more areas and more activities at its starting and end points. The video footage provided sufficient evidence to illustrate this fact: the activity that preceded the walking and the activity that succeeded the walking were recorded. For example, a supervisor may have been in the office working on a report, walked over to the hostess stand, walked guests over to their table, walked to talk to an employee, and then walked back to the office to do more paperwork. In total, the video footage showed that there were 348 distinct combinations of activities preceding and following the activity of an assistant manager or supervisor walking. The surveillance cameras also included footage for other employees. In contrast, there were fewer than 20 distinct combinations of activities preceding and following the activity for the hostess and the busing person.

This analysis clearly demonstrates that a supervisor performs walking activities at the worksite to maintain a smooth flow of operations. The large number of combinations of tasks also indicates that the direction of and tasks preceding and following the walking activity vary widely and require the supervisor to make judgment calls as to where to go and what to do based on his or her assessment of the situation. Implicitly, a supervisor has to assess what is necessary to keep operations running smoothly, prioritize the options, and then decide what to do. This has to be done in a matter of seconds, particularly when it is busy. Traveling/walking at the worksite facilitates this need for prompt decision making and therefore was classified as "exempt" based on the results of this observational study.

V. *Conclusion*

Observational studies in the context of wage and hour litigation are an important way to collect otherwise-unavailable data. For purposes of litigation, the data obtained through observational studies may be used to assess questions of liability and to potentially quantify damages to the class if liability is found. Observational data can also provide input to more elaborate models of damages and exposure. As in all situations in which data are incorporated into a broader model, conclusions are only as good as the input data—the old saying of "garbage in, garbage out" holds true.

This paper offers a framework for ensuring that data from an observational study will be reliable. The framework emphasizes the importance of making data quality a focal point, beginning with the design and planning phase of a study. Further, concurrent data quality-control procedures have to be part of the implementation process while an observational study is being conducted. Data quality simply cannot be an afterthought. A case study illustrated how to take proactive steps to achieve data quality in the planning and design phase, as well as during the implementation phase, so that results of such a study are reliable and defensible.

Dawn Eash
Berkeley Research Group

Dawn Eash consults on matters in the fields of labor and employment, intellectual property, commercial damages, public policy, healthcare, technology, and securities. Ms. Eash's experience includes sampling and survey design, advanced statistical and analytical methods, and construction and validation of complex databases.

Ms. Eash has conducted analyses and studies specifically relating to matters involving class certification; lost profits; wage and hour class actions; discriminatory hiring, promotion, and termination practices; breach of contract disputes; business valuation; quality control; government audits of insurance claim fraud; and securities fraud.

Contact Information:
Email: deash@brg-expert.com
Phone: 310.499.4823

Observational Studies in a Litigation Context: Important Practices in a Large-Scale Case Study

Abstract

Observational studies are a primary research method using direct observation to collect data on a population of interest. Observational studies are a common research method within academic research and have steadily gained acceptance in the courts, in particular in the area of labor and employment litigation. While the basic concept of observing and recording a subject's behavior to inform an expert's opinion on job duties and duration is straightforward, the actual implementation of a large-scale observational study is a complex logistical endeavor. This article anecdotally describes important practices Berkeley Research Group used in a large-scale observational study. Although each study presents its own set of anomalies and special circumstances, the practices discussed with respect to study design, execution, and post-study analysis should be carefully considered.

I. *Introduction*

Observational study, also referred to as direct observation, is a method of collecting primary data firsthand using an observer.[1] For example, if one wants to know what tasks are performed by high school English teachers (e.g., instruction, discipline, designing tests, administering tests, field trips, grading, correcting book reports), and how much time is spent at each task, one could either ask teachers to complete a questionnaire/survey or observe their activities directly. The advantage of direct observation (as compared to

1 E. Prien, L. Goodstein, J. Goodstein, and L. Gamble, "A Practical Guide to Job Analysis" (2009), pp. 28–29.

self-reports, interviews, or surveys) is that the data are collected firsthand in real time, reducing errors of memory and interviewer bias. Direct observation is most helpful with jobs involving physical activity, as an observer more easily ascertains physical tasks than intellectual tasks. Conversely, the disadvantage is the potential "audience effect" that occurs due to the presence of an observer recording tasks performed (i.e., the subject modifies his or her behavior when under observation).[2]

While the results of observational studies have been widely discussed in the academic literature, there has been little practical discussion of how large-scale observational studies should be designed, implemented, and administered.[3] The purpose of this article is to define and discuss best practices in the context of a large-scale observational study Berkeley Research Group conducted on behalf of a Fortune 500 company (the Company).

The Company's primary objective was to determine the amount of time sales representatives (reps) worked daily and weekly to calculate the potential exposure in an employee misclassification lawsuit.[4] A secondary motivation was to more thoroughly understand the tasks and duration of tasks performed by this type of employee to improve the efficiency and performance of their sales force.

2 Ibid.

3 In management literature, observational studies are referred to as "structured observation." See, for example, Jersild et al. (1939), Mintzberg (1968), Martinko and Gardner (1985), Stewart (1989), Tengblad (2006), Stewart (2007), and Mueller et al. (2012).

4 An employment misclassification case most commonly asserts that plaintiffs have been incorrectly categorized as exempt when the nature of their jobs is truly non-exempt and hence due unpaid overtime backpay. See Fair Labor Standards Act, "Coverage under the FLSA" (2003), accessed at: http://www.flsa.com/coverage.html

II. *Research Methods*

To determine the amount of time worked daily and weekly by a particular employee group, a variety of primary research methods can be considered, including self-reports, interviews, and surveys. Such methods are limited by the fact that the respondents provide opinions about their own work time that cannot be validated and could be incorrect (e.g., due to faulty memory or even intentional inflation). However, these methods are often more logistically simple to administer given that the subjects merely report activity summaries rather than being observed individually firsthand for the duration of each task completed. Therefore, researchers with time and/or cost constraints may prefer these research methods.

In our study, there was a possibility the subjects might report an increase in their tasks and inflate the time spent performing those tasks. Given this potential effect, the benefits of directly observing subjects outweighed the increased study cost and length associated with an observational study. The "audience effect" disadvantage was addressed directly in the study design and implementation, as detailed in the sections below.

III. *Study Design*

The study design is a crucial element of an observational study to ensure the results can be relied upon with confidence and key questions can be answered from the observed data. First, the researcher determines the sampling method and design (e.g., simple random sample or a convenience sample). Second, the researcher addresses data collection method and design (i.e., determine exactly how all time will be observed and recorded). Third, the researcher should consider and address potential shortcomings introduced by choosing a particular research method (e.g., counterbalancing the "audience effect" and accounting for time worked that cannot be observed by a third party), if possible.

Sampling Method and Design

The sampling method should be based upon the study objective and subject population characteristics. For example, the researcher should consider whether to use simple random sampling, stratified sampling, or a more complex two-stage sampling technique. In our case, we wanted to ensure several subgroups of the Company were adequately represented for further analysis, and thus a stratified sampling approach was determined to be appropriate.[5] "Stratification" in a statistical context refers to dividing the population of interest into homogeneous subpopulations (e.g., a population of teachers divided by the subjects they teach). By predefining strata to consist of these homogeneous subgroups, the researcher can predetermine important subgroup selection rates, and overall sampling error may also be reduced.

In our study, we ascertained that there may be significant variation in task duration between sales reps on different product lines. Second, we ascertained that the geographic region the sales rep operated out of may significantly affect task duration (e.g., rural versus industrial states). Therefore, we designed a stratified sampling method with two stratification factors proportionally weighted according to the current distribution of sales reps. Using these selection criteria and standard statistical procedures, we calculated the appropriate sample size.[6]

5 William. G. Cochran, *Sampling Techniques*, New York: Wiley and Sons (1977). Chapter 5, "Stratified Random Sampling": "If data of known precision are wanted for certain subdivisions of the population, it is advisable to treat each subdivision as a 'population' in its own right."

6 Sample size is an important aspect of sampling design, but since it is not the focus of this study we do not discuss the actual calculations here. For further information regarding sample size calculations, contact the author directly.

Data Collection: Method and Design

Data collection is an important aspect of the observational study, as one must ensure obtained data can be relied upon with confidence. The data being collected from various observers in various states should be consistently observed and correctly recorded. Therefore, in our study each randomly selected subject was individually observed or shadowed by a trained observer.

Each observer was equipped with a tablet preloaded with BRG-customized software containing a list of tasks most commonly performed by the sales reps. The task list was compiled from sales rep job descriptions and further refined by results obtained from pre-study results (discussed in more detail below). The software allowed an observer to simply tap the button associated with a particular task. For example, if an observer wanted to record a sales rep driving to his/her first appointment of the day, the observer would tap the button labeled "First Commute" under the "Driving" category. This tap activates the software that immediately begins recording time spent on "First Commute." Further, if the sales rep arrives at his/her first client site and must wait 10 minutes to see that client, the observer would tap the button "Check-In/Wait" under the "Sales Call" category. This tap triggers the software to stop the clock on "First Commute" and start the clock recording time spent on "Check-In/Wait." When the client becomes available for the meeting, the observer would tap "Discussion with Client," which stops the clock on "Check-In/Wait" and starts the clock recording "Discussion with Client." In addition, the software allows an observer to record notes while a task is being timed.[7]

7 If there was a gap in time (usually seconds) between stopping one task and starting the next, the start time of the second task was adjusted to be equal to the ending time of the first task.

The tasks recorded by each observer were transmitted continuously throughout the day to a secure server, allowing for real-time investigation of the results. The data was backed up and stored to preserve its integrity and facilitate ongoing data analysis. This process was continually monitored to ensure that complete observational data were uploaded and securely maintained.

Consideration of the "Audience Effect"

A key consideration in the design of the study concerned counterbalancing the potential "audience effect" caused by introducing a third-party observer into the subject's normal work atmosphere. Occasionally observational studies are performed by videotape as an easy way to eliminate an "audience effect"; however, in our case the sales reps were constantly on the move and therefore impossible to track via video. As mentioned earlier, since it was possible that the subjects might alter their ordinary activities to inflate their observed working hours, it was important to consider this potential "bias" within the study design to the extent possible. This was primarily addressed through the frequency and length of the observations.

Instead of one-day observations, each subject was observed for a full week (five working days) divided into consecutive and non-consecutive groups. Roughly half of the subjects were observed consecutively for five days, while the other half was observed intermittently throughout the study.[8] Repeated observations are expected to decrease the ability of the subject to consistently deviate from their normal routine. While one can deviate from their routine for an hour or perhaps a day, it would be much more difficult to deviate for an entire week. Therefore, extending the observational period with repeated observations decreased the likelihood of a bias effect.[9]

8 Some subjects were observed for four days owing to illness and could not be rescheduled prior to the end of the study. This was factored into the analysis.

9 See, for example, Cochran (1977): "when the same population… is sampled repeatedly, the sampler is in an ideal position to make realistic estimates" (p. 344).

On each day of observation, a trained observer accompanied a sales rep throughout the work day and, as noted above, used their tablet to record brief descriptions of the rep's activities and amount of time the rep spent carrying out those activities. The observer ceased his/her observations at the end of the rep's work day, as determined by whether the subject was returning home or had other personal obligations. If the subject was returning home, the observer was dropped off at the meet-up location, which was a public location as close as possible to the subject's home to give a realistic proxy of commute time. Otherwise, the subject and observer parted ways at the last sales location. This process was repeated throughout the observational study.

It is important to note that the subjects observed on consecutive days became aware that they would be observed and accompanied for a full week by observers when they received their first scheduling calls (which were made either on the Friday or Saturday preceding the first Monday observations). By contrast, subjects who were observed partially in one week and partially in subsequent weeks became aware of the study at or near the beginning of the study. With this additional lead time, these intermittently observed subjects might have been relatively more prone to inflate their work hours by booking additional meetings or scheduling distant sales calls. However, a post-study comparison of daily and weekly average work hours of continuously and non-continuously observed subjects found no statistically significant differences between these two groups.[10]

Accounting for Observed and Non-Observed Work Time

Due to the nature of these particular sales reps frequently traveling between client locations rather than operating in offices, it was necessary to individually

10 In fact, the fully observed consecutive subjects worked on average more hours than the fully observed intermittent subjects, although this difference is not statistically significant.

observe reps traveling to and from appointments. The amount of time reps spent performing work that could be directly observed was classified as "observed work time." Other tasks, such as completing and transmitting sales-related documents to the Company and reading and responding to email messages, could not be directly observed in this study. This work was often completed from home or hotel rooms during early morning and evening hours when it would have been both impractical and invasive to directly observe such work. To account for the amount of time spent on such work, the Company established a dedicated website. Sales reps were instructed to report their (non-observed) work at home or in hotel rooms to the website on a daily basis. This work is referred to hereafter as "website time."

The third category of sales rep work time that was accounted for and included in this study was "long commute time." This phrase refers to the amount of time of their work days that sales reps spent commuting (typically by driving) to their first appointments and from their last appointments to their homes. According to the Fair Labor Standards Act (FLSA), beginning and end of day commute time that exceeds one hour (i.e., 60 minutes) in either or both instances is compensable. Therefore, "long commute time" in this study refers to the amount of time beyond 60 minutes in the morning and 60 minutes in the evening that sales reps spent commuting.

To summarize, the daily and weekly compensable work time for the sample of sales reps that was included and reported in the study consisted of "observed work time," "website time," and established "long commute time."

IV. *Study Execution*

As described above, it is fundamental to align the study design with the study objectives and limitations. Perhaps of equal but sequentially secondary importance is the execution of those selected methods and designs. Once the

design is determined, the study should be executed with careful attention to detail, including but not limited to training observers, conducting a pre-study, scheduling subjects' observation periods, and troubleshooting.

Observer Training and Pre-Study

Trained, college-educated individuals served as the observers for this study. Prior to these observers being deployed in the field, a "pre-study" was conducted in which three sales reps were each observed for two days. The objective of the pre-study was to field test the study methodology and identify logistical or technical issues which could warrant further consideration. Shortly after these pre-study observations were completed, follow-up telephone calls were made to the observed reps. Reps were asked about their experiences being observed, including the observers' introductory comments and any work performance effects of being observed. This pre-study and follow-up process indicated that the observational method, task list, and manner in which work time was recorded were well suited to the observational study design and key objective. (The recording of sales rep work time is more fully discussed below.)

The pre-study and follow-up process also indicated that the training of observers needed to be slightly modified to include more detailed instructions about how observers should introduce themselves, describe their observer role, and conduct themselves during the observational process. In particular, observers were instructed to:

- Dress in business attire similar to their subjects
- Not initiate conversations with sales reps unless absolutely necessary to conduct observations
- Not engage in extended conversations or discuss non-business related matters with sales reps

- Not answer questions about what they recorded during the work day
- Not inform sales reps about the purpose of the study[11]

These instructions were issued in order to have the observers be as unobtrusive as possible so as to not influence the study findings.

Scheduling Observations and the Timing of When Sales Reps Were Informed

The scheduling of sales reps to be observed was a logistically complex process that began with the selection of the rep sample based on the aforementioned randomized stratification specification. There was sensitivity to the dissemination of information to subjects, and it was decided the reps' managers would be the first points of contact for the subjects selected for observation. BRG notified the Company of the identities of selected reps the week prior to the first day of observation, and this information was transmitted to the respective managers. BRG's professional staff then called or otherwise contacted each rep directly to schedule the specific days on which the reps would be observed. Those calls were made on Fridays, with the observations typically planned to begin the following Monday. In this way, the subjects of this study did not know that they were going to be observed until shortly before the observations actually began, and also did not know whether they would be observed for five consecutive or non-consecutive days until shortly before the observations actually began. This was an important factor in reducing the potential of the subjects altering their behaviors or schedules. This process continued during each week of the observational study until the end of the study.

11 The Company communicated the purpose of the study to sales reps. The Company instructed sales reps to immediately and directly report to their respective supervisors any issues that arose regarding sales rep–observer interaction.

Troubleshooting

Despite the attention given to all aspects of a study, anomalies will most likely still occur during the actual study. In the case of a large-scale nationwide observational study involving hundreds of subjects, observers, managers, BRG staff, and client staff, it is important to have a centralized "hot line" available 24/7 to respond immediately to last-minute questions, problems, or both.

V. ***Post-Observation Reconciliation***

The full observational study was conducted over a two-month period. At the conclusion of the study, the direct observation data was reviewed for context, the "website time" was reconciled with observed time, California rest breaks were accounted for, and lastly any compensable time unrecorded by the observer or website was imputed.

Direct Observation Review

The workday of a sales rep is complex and composed of granular tasks (e.g., waiting to see a client, reading a personal email message, responding to work-related email messages) that must be categorized according to their purpose. Some of these tasks are relatively easy to categorize (e.g., "Discussions with Client" is obviously "observed work time," and "Personal Appt" is obviously "personal time"), while other tasks must be considered in context (e.g., a personal phone call made while waiting to see a client is "observed work time" and not "personal time" because it was made while simultaneously performing a work-related activity).

For clarity, tasks that might be considered differently according to context (e.g., taking a phone call while waiting for a client versus placing a personal

phone call after arriving at the client site but before going inside) were explicitly categorized by occurrence. Accordingly, time spent by the sales reps on various activities was grouped into the following time categories: "observed work time," "website time," "long commute time," "personal time," and "commute time."[12]

Website Time

The time worked by sales reps at home or that could not otherwise be observed due to privacy or logistical constraints was an important component of reps' total compensable time. The website time was the best estimate of this unobserved time, but by definition this method is less precise than the time physically observed and recorded by an observer. Although the time reported by reps to the Company website could not be independently verified, numerous efforts were made to encourage sales reps to enter their unobserved time fully and accurately. These efforts included explicit email instructions, conversations with the managers, and in some cases extended deadlines for submitting unobserved work time. To the extent that website time conflicted with observed time, the observed time was given preference in the analysis.

California Rest Breaks

In order to account for paid rest breaks as mandated by the California Labor Code, each California sales rep was allocated 10 minutes of paid personal time (to the extent personal time, excluding personal meals, was taken) per four hours of work time.[13]

12 "Commute time" is commute time of less than one hour.

13 State of California Department of Industrial Relations, "Rest Periods/Lactation Accommodation," Division of Labor Standards Enforcement (revised March 4, 2011), accessed at: http://www.dir.ca.gov/dlse/faq_restperiods.htm

Imputed Time

The logistical requirements necessary to capture every task performed by the sales reps during several work days were complex and extensive. Consequently, some work time could not be directly captured and therefore needed to be imputed, including flights, commutes, study days not entered into the website, and days unable to be scheduled.

VI. ***Data Analysis***

The final resulting dataset detailed how each day was distributed amongst "observed work time," "website time," "long commute time," "personal time," "personal meal time," "total first commute time," "total end commute time," and "observer meet-up time." Rest-break time for California-based sales reps was then added to these data, yielding the adjusted total daily and weekly compensable time.

Results

From the more than 600 separate work days observed, the data yielded a number of important metrics related to tasks performed, duration of tasks, and even on which days certain tasks were generally performed.

Weighting the recorded data according to the current distribution of sales reps in each stratum, the estimated overall average weekly work time for the current universe of reps was slightly above 40 hours with a standard error of approximately 45 minutes. Individual sales rep weekly compensable time ranged between a low of approximately 30 hours and a high of 60 hours. While the range of individual sales rep total weekly compensable time was quite wide, the standard error was relatively small.

The first four days of the week were largely similar, with average daily work times of approximately 8 hours and 15 minutes; however, not surprisingly, the

daily average was significantly lower on Fridays at approximately 7 hours. The minimum hours worked during any single weekday was 1 hour, and the maximum was 16 hours.

Non-California Overtime Calculations

Sales reps working outside of the state of California are subject to the FLSA definition of overtime, which specifies that any time worked above the threshold of 40 hours in a work week is eligible for overtime pay. Accordingly, if sales reps worked fewer than 40 hours in a week, they are not entitled to overtime pay, and their calculated overtime is equal to 0.[14] Weighting the recorded data according to the current distribution of sales reps in each stratum, the estimated overall average weekly overtime for the current universe of non-California sales reps was approximately 2 hours with an overall standard error of 30 minutes.

California Overtime Calculations

Sales reps working in the state of California are subject to the California Department of Industrial Relations (CA DIR) definition of overtime, which specifies that hours worked over 8 hours in a workday and over 40 hours in a work week are eligible for overtime pay.[15] Accordingly, if sales reps worked fewer than 8 hours every day and fewer than 40 hours in a week, they are not entitled to overtime pay, and their calculated overtime is equal to 0. Weighting the recorded data according to the current distribution of sales reps in each stratum, the estimated overall average weekly overtime time for the current

14 Mathematically, the overtime variable can be expressed and calculated as the maximum of 0 and the difference between compensable hours and 40.

15 State of California Department of Industrial Relations, "Overtime," Division of Labor Standards Enforcement (2013), accessed at: http://www.dir.ca.gov/dlse/faq_overtime.htm

universe of California sales reps was roughly 4 hours with a standard error of 35 minutes.

VII. *Conclusion*

This article presented a large-scale observational case study in an effort to highlight important practices related to research method, study design and execution, post-observation reconciliation, and data analysis. We note that case-specific facts will guide design of a large-scale observational study and that these guidelines are not necessarily universal. Nevertheless, consideration of these practices will safeguard against potential weaknesses of observational studies and increase the validity, integrity, and accuracy of the results. These practices can be summarized as follows:

1. Observational study is a powerful research method for gathering primary data, but the study design should attempt to counteract the "audience effect" as much as possible (e.g., repeated observations).

2. To the extent there are important subpopulations within the population of interest, consider a more versatile sampling method (e.g., stratification).

3. Observations should be recorded and tracked by a device with secure real-time capabilities to ensure data integrity and facilitate real-time data analysis.

4. A pre-study should be performed to ensure that observer training is comprehensive and no details have been overlooked.

5. Observer training should be performed after the pre-study to ensure that observers are properly equipped to handle introductions and issues that may arise during the study.

6. Task lists should be compiled based on job descriptions and discussions with the client, and further refined if needed based on the results of the pre-study.

7. Observers should follow a strict dress code similar to the subjects of interest.

8. When scheduling observations, subjects should be given as small of a lead time as possible to decrease the opportunity to make changes in their behavior.

9. At the home office, a dedicated agent should be provided to whom subjects and observers can direct questions and problems for troubleshooting. This also allows the home office to disseminate potential problems and appropriate solutions to other staff on an ongoing basis.

10. Post-study observed time should be reviewed and classified according to context.

11. Alternative methods, such as a website, should be utilized to account for unobserved compensable time.

12. Unrecorded time should be imputed into compensable time.

13. All time recorded by the observer should be reconciled with secondary sources, including website time, observer notes, and other available resources.

About Berkeley Research Group

Berkeley Research Group, LLC is a leading global expert services and consulting firm that provides independent expert testimony, litigation and regulatory support, authoritative studies, strategic advice, and document and data analytics to major law firms, Fortune 500 corporations, government agencies, and regulatory bodies around the world.

From testifying in high-stakes litigation to consulting on large-scale projects, BRG experts and consultants combine intellectual rigor with practical, real-world experience and an in-depth understanding of industries and markets. Their expertise spans economics and finance, data analytics and statistics, and public policy in many of the major sectors of our economy, including healthcare, banking, information technology, energy, construction, and real estate.

Berkeley Research Group is headquartered in Emeryville, California, with 24 offices across the United States and in Australia, Canada, Latin America, and London, United Kingdom.

For personalized assistance in locating an expert, please contact Client Services at 1.877.696.0391 or email us at info@brg-expert.com.

www.brg-expert.com