American Journal of Infection Control 42 (2014) 1067-73

Contents lists available at ScienceDirect

American Journal of Infection Control

journal homepage: www.ajicjournal.org

Major article

Observer accuracy and behavior analysis: Data collection procedures on hand hygiene compliance in a neurovascular unit

Krista L. Hinz MA [a,*], Heather M. McGee PhD [a], Bradley E. Huitema PhD [a], Alyce M. Dickinson PhD [a], Richard A. Van Enk PhD [b]

[a] Department of Psychology, Western Michigan University, Kalamazoo, MI
[b] Department of Infection Prevention and Epidemiology, Bronson Methodist Hospital, Kalamazoo, MI

Key Words:
Behavioral science
Data accuracy
Response effort
Accurate reporting
Data collection integrity

Background: Although observational studies are popular, little has been done to study the integrity of human observers and the data collection process. Issues of data collection integrity threaten functional findings, leading to problematic interpretation and decreased replication. In our study the response effort associated with hand hygiene data collection in a hospital setting was manipulated using an altered data collection tool.

Methods: A counterbalanced ABAB design was implemented across 2 semesters of a hand hygiene data collection practicum course.

Results: When response effort increased, compliant audits decreased and when response effort decreased, compliant audits increased. These results were statistically significant, with an overall level change *z* that had a *P* value of .001 (first semester) and .007 (second semester).

Conclusion: These findings may warrant an increased awareness of data collection procedures where recording options include a less effortful response. The results of our study support basic research on response effort and choice behavior in an applied setting, bringing into question the integrity of data collection procedures and the integrity of the data collected. These results also suggest the need for standardizing reporting systems to ensure hand hygiene collection and reporting procedures are comparable across settings.

Isolating hand hygiene as a causal variable for hospital-acquired infections (HAIs) was demonstrated in the late 1840s when Ignaz Semmelweis was able to garner empirical support for the concept of attributing the transmission of puerperal fever to the unclean hands of health care workers (HCWs).[1] Hand hygiene was, and still is, the most important practice for preventing the transmission of HAIs. However, despite all we have learned through research and technological contributions since transfer of disease was first postulated, hand hygiene compliance still remains a problem.

During 2002, an estimated 1.7 million patients in the United States acquired an HAI and, of those, an estimated 99,000 patients died as a result of the infections. This situates HAIs as the fifth leading cause of death in US acute care hospitals.[2] Research validates that HAIs decrease as compliance to hand hygiene protocol increases.[3] Research further suggests infection rates can be decreased by 33% with compliance to hand hygiene protocols.[4] Although hospitals have long had policies requiring HCWs to conduct hand hygiene between patients, reported compliance rarely exceeds 50%.[5]

HAND HYGIENE PROGRAMS

Most hospitals are now implementing programs to measure and improve hand hygiene compliance. In addition to the social significance of increased quality of care and safety to patients and HCWs, reducing HAIs decreases financial loss for organizations. During 2007 additional treatments and longer hospital stays resulting from HAIs were responsible for an estimated $35.7 billion to $45 billion in extra health care costs.[6] As of October 2008, the Centers for Medicare and Medicaid Services enacted a revision to payment systems that excludes coverage of HAIs. The revised system further prevents health care organizations from passing additional cost of HAIs to patients.[7] Many private insurers

* Address correspondence to Krista L. Hinz, MA, Department of Psychology, Western Michigan University, 64233 Baseline Rd, South Haven, MI 49090.
E-mail address: kristalyn79@yahoo.com (K.L. Hinz).

Conflicts of interest: None to report.

0196-6553/$36.00 - Copyright © 2014 by the Association for Professionals in Infection Control and Epidemiology, Inc. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ajic.2014.06.017

are following suit, leaving the burden of cost to health care organizations.[8]

Unfortunately, most interventions on hand hygiene compliance have short-lived success. To be effective, a program must become part of a permanent practice. Furthermore, to obtain long- or short-term success, the program must have administrative support. Providing a bolster of administrative support aids the intervention program by ensuring that change will take effect via consequential action. The naturally occurring consequences for engaging in hand hygiene within a hospital are often punishing and ineffective. To follow suggested protocol, staff must use hand hygiene measures frequently. To do so requires response effort, interruptions in routine, and time away from patients or other tasks. In addition, frequent hand hygiene increases dry and chapped hands, which is physically aversive. These consequences are immediate and punishing. Escape from negative covert verbal behaviors surrounding perceptions of infection and personal hygiene (eg, the worker "feels" dirty) may provide some reinforcement, but the probability of microbial transmission is often perceived as unlikely.[9] That is, they may engage in hand hygiene to "feel clean" if a patient encounter made them "feel dirty." Furthermore, unlike a medication error, it is unlikely that the consequence of patient harm from contaminated hands will be connected back to the HCW responsible. Without this feedback, crucial negative covert verbal behaviors (eg, "I am harming my patient by not washing my hands") are unlikely to occur. Additionally, the existing environmental contingencies from the organization—social approval or disapproval—are not probable or valuable enough to control the behavior. To evoke change, organization-wide consequences need to be established that will support hand hygiene behaviors. To ensure that they are enacted, high levels of administrative support must be employed.

Another problem with current hand hygiene programs is the lack of uniformity in the dissemination of results due to varied data collection procedures and methodology. Operational definitions of what does and does not constitute a hand hygiene opportunity differ, as do the data collection methods. In addition to varied methods and criteria used, reported research does not convey specifics of their components.[3] This may contribute to difficulties with operational definitions and discernment of auditing opportunities.

Direct observation of hand hygiene behavior allows us to see hand hygiene behavior as it is occurring.[10] Despite this benefit, there are drawbacks to direct observation, including resource allocation requirements, a lack of universal standardization of the auditing process, reactivity, differences in training and experience of the observers, and differing operational definitions between organizations. Research on the use of human observers cautions researchers to ensure the observers are conducting accurate observations. Factors such as reactivity, observer drift, the recording procedure, reliability, complexity and demands of the task, subject, and setting may all compromise the data collection practice.[11] Furthermore, using preexisting staff to audit the hand hygiene behavior of other employees introduces the potential for biased data. Falsification of data may result from negative treatment by peers, pressure from the organization to do well, and/or punishing consequences that may fall on a particular department or the organization as a whole for results that do not meet a set goal. Falsification and/or withholding data prevents an accurate representation of an organization's hand hygiene behavior and fails to identify areas that need improvement, although adding inter-rater reliability procedures can help to reduce this to some extent. This said, studies have demonstrated that direct peer observation can be effective in reducing accidents, even if data are not always perfectly reliable.[12] However, accurate data is essential for the integrity of research findings and it is certainly essential with regard to disease surveillance and hand hygiene. Accurate data collection lends awareness to problem areas that may, in turn, lead to optimization of processes that improve hand hygiene. Additionally, when dealing with human observers and error, tight control in data collection and methodology is essential for accurate representation of compliance. Inconsistent and inaccurate measurement of adherence results in reports that are questionable, making comparisons of organizational compliance between institutions difficult.

Little research has been conducted on the integrity of human observers and the data collection process, aside from employing inter-rater reliability (IRR) procedures. Energy is typically focused on treatment integrity by making sure the intervention was implemented as planned. Even then, the assessment of treatment integrity has been relatively low.[13] Antecedent tools in the form of instruction and standard operating procedures (SOPs) get desired adherence started, but consequences are necessary for maintenance. Although there may be set expectations and clear operating procedures in place for both the auditors and the employees being audited, behavioral research confirms that what really controls behaviors are the consequences in place. When there is a lack of consequence for employees engaging in hand hygiene and also for auditors documenting the behavior, or if the natural consequences support the wrong behavior, you will get undesired behavior that SOPs cannot fix. Integrity of both the independent and dependent variables is essential and incomplete analysis of their respective integrity threatens functional findings, leading to problematic interpretation and decreased replication.

In the science of human behavior—behavior analysis—response effort refers to the amount of effort required to complete a task. That is, how much effort is needed to accomplish a particular behavior. As response effort increases and a task becomes increasingly difficult, there will be an effect on an individual's behavior. Basic experimental research has shown that if an organism is presented with 2 choices that have the same outcome, but with differing response efforts, the organism will allocate more time to the less effortful response.[14] That is, when given an easy response versus a harder response, the organism will choose the easier response. Research further demonstrates that response rates decrease as response effort increases.[15] There is little applied human research investigating the effects of response effort on responding when given a choice between 2 responses.

Differing response efforts in data collection is not uncommon. For example, when collecting data on the occurrence of a behavior, negative or positive, there is response effort for collecting data on the target behavior—but little or no response effort for collecting data on the absence of the behavior. The occurrence of the behavior may require the observer to record times, dates, settings, antecedents, and consequences. These constitute higher response effort because they require additional time and attention on the part of the auditor. Further, the collection of longitudinal data may be subject to increased fatigue and lax practices by the observer. To promote optimal data collection, response dimensions, specifically response effort, should be equal across all levels of responding. Because experimental data have shown that organisms prefer low effort responses, equalizing effort should prompt auditors to choose the correct response, instead of the easiest. When response effort differs between 2 choices, every effort should be made to decrease the response effort associated with the more difficult task.[16] However, this is not always possible.

Our study

Although SOPs were in place for hand hygiene and the data collection process, the high percentage of reported compliance was concerning when compared with national averages. The

purpose of our study was to discern the validity of the data collection process with regard to the response effort involved with collecting hand hygiene data. This included assessing if increasing the response effort associated with marking compliance to match that associated with marking noncompliance would improve the accuracy of data collection. That is, equalizing the response effort, making it just as hard to mark compliance as it is to mark noncompliance. In our study, the effort associated with recording hygiene compliance data was lower than that required to record noncompliance data. To document compliance, the auditor simply had to see an HCW engage in hand hygiene as he or she entered or left a room. To document noncompliance, the auditor had to additionally mark what was touched in the room resulting in the noncompliant audit. This involved active involvement with auditing. It was easier to document compliance because no vigilance was required. However, to document noncompliance required the auditor to be vigilant during the process. He or she had to physically move and pay attention to what was touched in the room. When equaling the response effort for compliance to match the response effort of documenting noncompliance, vigilance was necessary for documenting compliance and noncompliance. In this setting, it was not practical to decrease the response effort associated with recording noncompliant data. Therefore, an increase in response effort associated with recording compliant data was manipulated. The study was conducted with external, nonbiased auditors.

METHODS

Participants and setting

Participants in our study were undergraduate psychology students at a midwestern university who were enrolled in an undergraduate Industrial/Organizational practicum that took place at a local hospital. Using students to audit data is not uncommon.[17] The use of external auditors removes bias that can lead to inaccurate reporting. The hand hygiene program at the hospital was run by the first author (KLH), who was also the instructor for the practicum. That is, the hand hygiene program, run through the practicum, was part of normal business practice. In compliance with the university's policy, the hospital's policy, and federal regulations, human subjects institutional review board permission was obtained from both the university and the hospital to ensure the protection of research participants. The auditors were blind to the purpose of the study. A total of 7 undergraduate student auditors across 2 semesters participated in the study with informed consent. The auditors directly observed patient care units at the hospital, documenting compliance and noncompliance of the employees using a paper audit tool. The setting was a 404-bed hospital located in the Midwest, encompassing 23 acres. The hospital consisted of a north and south campus comprising 22 units with a total of 404 private rooms.

Apparatus and materials

Hand hygiene auditing tool

SOPs were already in place for hand hygiene and the data collection process. The hand hygiene auditing tool was used to collect data on compliance. The back of the audit tool doubled as a job aid and denoted which units to audit, which rooms were included in each unit, the types of employees audited, and a description of the color-coded scrubs worn by the differing types of employees. The auditing tools were modified slightly during the intervention phases as a visual reminder of changes in the auditing process.

Experimental design

An ABAB and BABA within-subject design was used. During the first semester of auditors, an ABAB design, where A = existing response requirements (less response effort for documenting compliant behavior and more response effort for documenting noncompliant behavior) and B = equal response effort requirements (response effort that is equal for documenting compliant and noncompliant behavior) was used to examine the effects of raising the response effort associated with collecting data on hand hygiene compliance. During the second semester of auditors a BABA design was employed. The phase changes were implemented every 3 weeks.

Because of the limited number of auditors, it was not feasible to use a between-group study where a large number of participants serve as a control for between-participant variability. Instead, it was more appropriate to use a within-subject design that permitted use of a smaller number of participants by which each auditor served as his or her own control. That is, an auditor's data collection behavior in 1 phase was compared against his or her own performance in an intervention phase, removing the need to control for between-auditor variability.[18,19] Given the practical constraints inherent in the hospital's practicum-based hand hygiene program (ie, few hand hygiene auditors), we chose to examine the performance of a smaller number of individuals over multiple trials to assess the effects of repeated exposure to the independent variables, believing that performance may differ with repeated exposure rather than 1 or 2 exposures as is typically done in between-group studies.

Procedures

Each auditor collected 10 hours of data per week following the general practicum procedures. Because of patient privacy, auditors were not allowed to enter rooms and instead audited from the hallway in the event they had a clear view of the room and HCWs. The auditors collected data from all 22 units of the hospital, but data from only 1 controlled unit (the neurovascular unit) was used. The manager in this unit agreed not to provide the staff with consequences in the form of performance feedback and evaluative statements during the duration of the study. The specific data collected varied based on the research condition in effect. The hospital's hand hygiene policy clearly specified when a HCW should perform hand hygiene during a patient encounter. At a minimum, hand hygiene should have been initiated upon entry, before an HCW touched a patient or a patient's environment, then again before and after gloving, and finally before the HCW left the room.

Dependent variable

The dependent variable was a measurement of the percentage of opportunities for hand hygiene marked as compliant by the auditors. The data were obtained from the data collection sheets completed by the auditors. Compliance was calculated as the number of compliant audits divided by the total number of compliant and noncompliant audits.

Baseline

During each A phase, auditors adhered to the hand hygiene data collection practices that were in place. Each baseline collection phase lasted 3 weeks. During the baseline phases, the auditors noted pertinent information regarding the observations: date, time, unit, employee type, if the opportunity took place before or after an

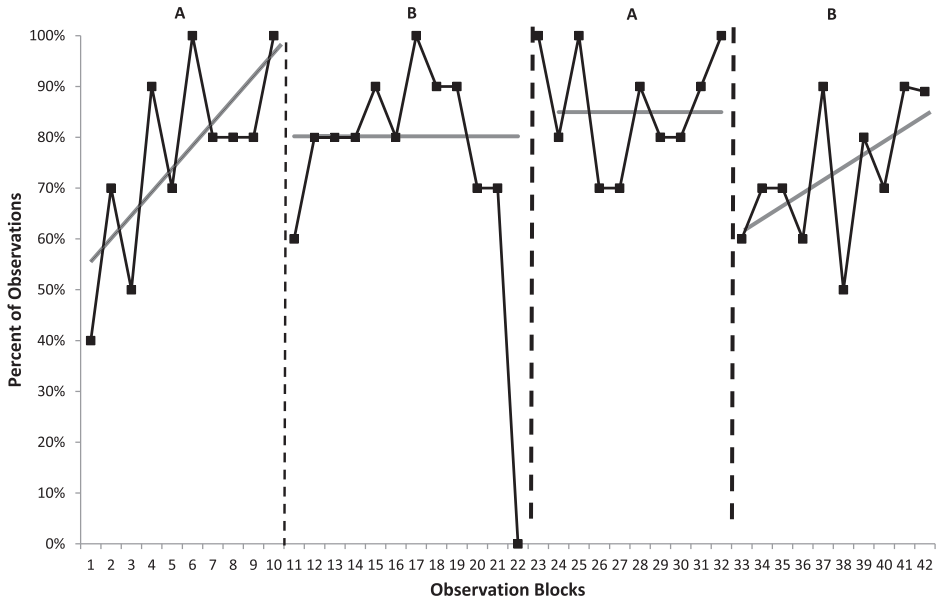

Fig 1. Percent compliance across phases during the first semester.

HCW contacted a patient's environment, if the HCW was compliant or noncompliant, the room number, and if hand hygiene occurred inside or outside of the room. In addition, if an HCW was noncompliant, the auditor had to observe and document what the HCW touched. They did not need to observe and document what was touched for compliance.

Independent variable

During each B phase, the auditors collected compliance data as they had during baseline, with the additional requirement that they indicate what, if anything, the HCW touched while in the room for both compliant and noncompliant auditing. During baseline, this step was only required for noncompliant audits. As it was during phase A, it required more response effort to document noncompliance than it did to mark compliance. To mark compliance, the auditors had to observe an HCW engaging in proper hand hygiene behavior. To mark noncompliance, the auditors were required to observe an HCW physically contacting a patient environment without engaging in hand hygiene and, in addition, record what the HCW touched in the patient environment. This equalized the response effort of both choices, removing the possibility of a less effortful response. The intervention phases lasted 3 weeks. The changes in the requirements of data collection were described to the students during weekly meetings at phase changes. At this time, the researcher collected any auditing tools the auditors had from the previous stage and distributed the phase-appropriate tools for the next condition.

The SOPs were developed in accordance with the World Health Organization 5 Moments of Hand Hygiene.[20] An opportunity to audit occurred when an HCW entered a patient room, examination room, or procedure room and physically contacted (touched) the patient or the patient's environment. A patient environment refers to anything within a patient's room. Objects and equipment inside a patient's room were audited as the transfer of microorganism to inanimate objects pose an infection risk to both patients and HCWs. In addition, other opportunities to audit occurred when an HCW exited a patient environment, put on gloves, or removed gloves. A patient encounter was not counted if an HCW walked into a patient's room and did not come into contact with a patient or environment.

IRR

To ensure the validity and reliability of the data with regard to homogeneity and consistency, IRR was calculated by dividing the number of times independent auditors agreed on the outcome of an observation by the total number of times the independent auditors agreed and disagreed. The auditors collected IRR data with each other or a research assistant for 35% of the observation sessions and obtained 98% agreement. In addition to assessing the validity and reliability of the data collection process, IRR was conducted on the measure of the proportion of total hand hygiene compliance to ensure that the auditors calculated IRR correctly. IRR on these calculations was completed by the first author (KLH), along with a research assistant, for 35% of all observation samples, obtaining 99% agreement.

RESULTS

The total number of observations in the neurovascular unit for the duration of the study was 555. Over the course of the first semester there were 4 auditors who conducted a total of 394 audits with A = 91, B = 111, A = 93, and B = 99 audits across the phases. There were 3 auditors during the second semester who conducted a total of 161 audits with B = 30, A = 53, B = 30, and A = 48 audits during the phases.

First semester

The purpose of the statistical analysis presented here is to evaluate the strength of agreement between the predicted outcome of the experiment and the actual outcome. The observed data are presented in Figure 1. The predicted outcome pattern for the first semester was that compliance percentage would be higher during low effort phases than during high effort phases. Specifically, the level would be relatively high during first phase, it would decrease during the second phase, it would increase during the third phase, and it would again decrease during the fourth phase. Predictions 2 through 4 are predictions of the direction of change; these predictions can be summarized as the following sequence of negative and positive signs: −, +, and −.

Table 1
Intervention analysis summary for first semester data

Type of change measure	Effect coefficient	*t*	*P*-value	95% Confidence interval	Standard effect size	*PS (OS)|z|**
Phase 1 trend	5.00	2.86	.01	—	—	—
Level change 1 (1,2)	−25.00	−2.12	.04	(−48.89 to −1.11)	−1.57	2.05
Level change 2 (2,3)	10.00	1.47	.15	(−3.79 to 23.79)	0.63	1.44
Level change 3 (3,4)	−25.42	−2.40	.02	(−46.92 to −3.92)	−1.60	2.29
Phase 4 trend	2.71	1.55	.13	—		—

PS/OS, predicted signs and observed signs.
NOTE. Overall level change = 16.50; overall level change $z = 3.33$ ($P = .001$); and robust $R^2 = .26$; and standard error of estimate = 252.
*$\sum_{c=1}^{C} PS(OS)|z_c| = 5.78$.

Preliminary analysis

A preliminary analysis was performed to identify the most appropriate model for the observed data. This analysis indicated that the model requires a slope parameter to describe the trend in the first phase, level change parameters for the changes from phases 1 to 2, 2 to 3, and 3 to 4, and a slope parameter to describe the trend in phase 4. The parameters of this model were fitted using methods described by Huitema.[21] Diagnostic tests were then carried out to evaluate independence of the errors and conformity with other model assumptions (eg, homogeneity of variances and normality of errors). The data conformed to the assumptions of the model in most respects; however, an outlier was present in the second phase. Therefore, the final estimation of the model was carried out using a regression procedure based on the robust general linear model that is described thoroughly in Hettsmansperger and McKean.[22] This estimation approach was used because it provides appropriate effect estimates even in the presence of outliers.

Main analysis

It can be seen in Table 1 that the chosen model includes 5 coefficients to describe behavior throughout the duration of the experiment. The first coefficient describes the increasing trend in the first phase. It indicates that the average percentage of compliant responses increased 5 points per block for the first 10 blocks. This suggests a learning effect within this phase. The second coefficient describes a 25-point decrease in level after the introduction of the high response effort condition. The third coefficient indicates that the level increased 10 points after the return to the low response effort condition. The fourth coefficient describes the decrease (25.42 percentage points) in the phase-4 level (relative to the phase-3 level) that occurred when the high response effort condition was reintroduced during phase 4. The last coefficient describes the average increase in compliance (2.71 percentage points per block) for the 10 blocks within phase 4.

The 95% confidence intervals for the level change coefficients and the standardized effect sizes are also shown in Table 1. The standardization of the level changes was based on the standard error of estimate, which is the standard deviation of the pooled within phase residuals. Hence, the standardized effect sizes represent the change in phase level expressed in standard error of estimate units.

Table 1 also provides a quantity that is required in computing the overall level-change test. The overall level-change test shown at the bottom of the table evaluates the strength of the argument that the predicted pattern of level change is consistent with the observed pattern. It incorporates information regarding both the agreement of the signs associated with the predicted level changes with the signs of the observed level changes, and the strength of the evidence for each individual level change.

There is perfect agreement between the predicted signs and the signs actually observed in Table 1 for the level change coefficients. It was predicted that the 2 changes from low to high response effort would lead to decreases in compliance; notice that the signs for the observed level change 1 and level change 3 coefficients are negative. Similarly, it was predicted that the change from high response effort to low would lead to an increase in compliance; the positive sign on LC2 indicates that this occurred. Using the formula below Table 1, these 3 agreements between predicted and observed outcome and the strength of the evidence associated with them (ie, the *P* values associated with the coefficients) were cumulated in the overall level-change test statistic *z*. The *P* value ($P = .001$) associated with this statistic leads to the conclusion that the evidence supporting the predicted effects of the condition changes is very strong. That is, there is strong evidence that when response effort increased, it led to a decrease in compliance auditing behavior and when response effort decreased, compliance auditing behavior increased. When recording compliance data became more difficult for the auditors, compliance data was recorded less often.

Again using the formula below Table 1, the proportion of the total sample variation that can be explained by the interventions is provided by the R^2 coefficient which is equal to 0.26. A value of this magnitude may be considered a large effect size. That is, more than one-quarter of the total variation in the experiment can be attributed to the independent variable manipulations.

Second semester

The second semester design consisted of 4 phases arranged in a BABA exposure sequence (unlike the first semester design that used an ABAB sequence). The analysis approach was essentially the same as was used for the first semester data.

Preliminary analysis

The model identified for the second semester data is somewhat simpler than the model for the first semester data. Because no significant slope was identified within the phases, the only change parameters required in the model are for level change. Hence, an intercept and 3-level change parameters are included in the model. All diagnostic tests of the adequacy of this model were acceptable.

Main analysis

Figure 2 illustrates the second semester data. The level change estimates and the associated inferential statistics are shown in Table 2. Before the data were collected it was predicted that performance would be higher during low effort phases than during high effort phases. Therefore, the signs associated with the three adjacent phase change predictions were: +, −, and +. Notice that the observed signs associated with the 3 level-change coefficients listed in Table 2 agree with the predicted signs. Although none of the individual level change coefficients has a *P* value $< .05$, the information cumulated from these individual tests yields an overall level change *z* that has a *P* value = .007. Hence, the cumulative data from all phases lead to the conclusion that the intervention clearly had the predicted effect pattern. As was found in the first semester study, high response effort was associated with low compliance and low effort was associated with high compliance. The R^2 coefficient presented at the end of Table 2 indicates that more than 40%

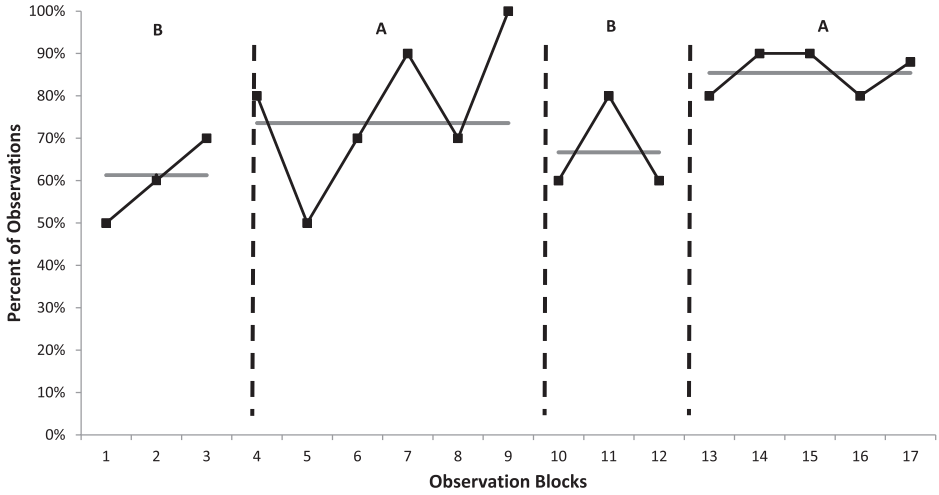

Fig 2. Percent compliance across phases during the second semester.

Table 2
Intervention analysis summary for second semester data

Type of measure	Coefficient	*t*	*P*-value	95% Confidence interval	Standard effect size	*PS (OS)|z|**
Phase 1 level	60.00	—	—	—	—	—
Level change 1 (1, 2)	16.67	1.85	.09	(−2.78 to 36.12)	1.31	1.711
Level change 2 (2, 3)	−10.00	−1.11	.29	(−29.45 to 9.45)	−0.79	1.065
Level change 3 (3, 4)	18.93	2.04	.06	(−1.15 to 39.02)	1.49	1.859

PS/OS, predicted signs and observed signs.
NOTE. Overall level change = 15.12; overall level change $z = 2.68$ ($P = .007$); $R^2 = .41$; and standard error of estimate = 162.
*$\sum_{c=1}^{C} PS(OS)|z_c| = 4.635$.

of the total sample variation was associated with the manipulation of response effort.

Comparison of first and second semester results

There are several ways to evaluate the similarities and differences between the results found for the 2 studies. The similarities are striking. The overall level-change was similar in both studies and both of them revealed clear evidence of intervention effects of the predicted form. The *P* values for overall level-change in the 2 studies were .001 and .007; a test on the difference between these 2 *P* values yields a *P* value = .64. Hence, even though the sample size was larger in the first semester than in the second, the overall evidence for a treatment effect is not significantly different in these 2 studies.

Because the 2 studies were designed to have the order of the condition manipulations reversed (ie, the first semester order was ABAB and the second order was BABA) it was possible to examine possible sequence effects. The standardized level changes associated with the 2 semesters provide a simple basis for evaluating this possibility. The average standardized level changes for the AB and BA sequences were similar (−1.58 for the AB sequence [first semester] and 1.40 for the BA sequence [second semester]). That is, the average standardized decrease associated with introducing the more difficult condition after the simpler condition was similar to the average standardized increase associated with introducing the simpler condition after the more difficult condition. Correspondingly, the single BA change (from phase 2 to phase 3) in the first semester increased the outcome score 0.63 standardized units and the single AB change (from phase 2 to phase 3) in the second semester decreased the outcome score 0.79 standardized units. (The raw AB and BA absolute percentage changes in the 2 semesters were identical.) The overall conclusion of these comparisons is that evidence for sequence effects is not strong.

An additional metric for comparing the 2 studies is the proportion of variation explained by the independent variable manipulations. The R^2 values were 0.26 and 0.41 for the first and second semesters, respectively. Most of the difference between these coefficients can be explained by the fact that the first semester data were more variable within phases than were the second semester data. This can be seen in Figures 1 and 2. Similarly, it is revealed in the pooled within-phase error variance estimates of 252 and 162 for the first and second semesters, respectively.

In summary, the results of the 2 studies are quite similar. The evidence for the existence of treatment effects within each study as well as the evidence for the consistency of these effects is strong.

DISCUSSION

Results of our study demonstrate that having differing levels of response effort involved with data collection affects data collecting behavior. The consequences for collecting compliant behavior and noncompliant behavior were equaled, removing an easier response that positively skewed the reported hand hygiene data. It was predicted, tested, and shown that changes from low to high response effort would lead to a decrease in compliance and the change from high response effort to low would lead to an increase in compliance. During the first semester, there was perfect agreement between the predicted signs during all 3 level changes denoting that the predictions were correct. Further, a very conservative *P* value ($P = .001$) elucidates that the evidence surrounding all 3 level changes is very strong. During the second semester, the level changes yielded a *P* value that was significant ($P = .05$), but not as conservative as the first semester. Cumulatively the individual tests led to an overall level change *z* that had a more significant *P* value ($P = .007$). The statistical findings of both semesters lead to the conclusion that the intervention markedly had the predicted effect pattern. When response effort increased, compliant audits

decreased. When response effort decreased, compliant audits increased. That is, when the response effort for both conditions was equalized, and there was no longer an "easier" compliance option, the auditors chose this option less often.

An additional statistical analysis was run to evaluate the differences between data collected during the first semester and data collected during the second semester. The results show that even though the sample size was larger during the first semester, there was no significant difference in treatment effect. Further, comparison of the average standardized level change between phases and semesters draws us to the conclusion that the evidence for sequence effects is not strong. There is a strong similarity between the studies with regard to clear treatment effect and consistency of effect.

These results can guide future research and practice by bringing to attention data collection processes. Comparing compliance between institutions is futile if accurate information is not being reported. Further, inaccurate data fail to identify compliance problems, placing patients at increased risk for HAIs.

Limitations and future research

One limitation to our study was the sample size of the observations. During the second semester, there were considerably fewer observations and, therefore, fewer observation blocks. This may have influenced the individual level change results. That is, whereas we saw significant level changes overall during the first semester data, we did not see significant individual level change. In future research, it would be beneficial if the number of observations were more consistent across semesters.

Another limitation of our study was the use of students as auditors involved in a practicum class. Each semester, we worked with new practicum students, which meant that student performers were still likely learning when data collection began, affecting our first phase each semester. Because of the semester-imposed time restrictions, we were only able to collect 13 weeks of data for each group, allocating approximately 3 weeks to each phase. Future researchers should consider working with staff that has stable positions, allowing performance to stabilize before intervening. Alternatively, if there are semester-imposed restraints, future researchers should consider using an ABA or a BAB design to allow stabilization of performance.

Finally, future research should explore a nonvariable environment. Because it was a hospital setting, the environment was constantly changing. A nonvariable environment, in addition to the absence of reactivity, would give a more precise account of human data collecting behavior. Conducting similar research in a controlled lab setting may result in bolstered effects.

These findings warrant an increased awareness of data collection procedures where recording options include a less effortful response. Basic experimental research has demonstrated that when presented with 2 choices, 1 being less effortful, an organism will choose the response with the least amount of effort.[14,23] The results of our study support these research findings in an applied setting, bringing into question the integrity of data collection procedures and the integrity of the data collected.

The results of our study also suggest the need for standardizing reporting systems to ensure hand hygiene reporting procedures are comparable across settings. Ensuring that the data collection and reporting of hand hygiene compliance between institutions are the same for comparison purposes is not only important for benchmark data used by organizations, but also for comparisons within research. Further, specification of the auditing tools used in data collection is essential to ensure that there are no unseen barriers to correctly reporting compliance.

SOPs in the form of education and instruction may get a behavior started, but consequences control the maintenance of HCWs engaging in hand hygiene and auditors accurately capturing hand hygiene behavior. To ensure desired behavior is enacted to protect HCWs, patients, and to also accurately report compliance, focus needs to be placed on the root of human behavior to discern why compliant behavior on either end is not taking place. Failure to do so breeds an assumption that HCWs and auditors will do the "right thing," even if it is harder and aversive to do so. This violates basic laws of human behavioral science. To promote organization-wide change for the health of our patients, in addition to accurate data representation, consequences need to be addressed.

References

1. Miller PJ. Semmelweis. Infect Control 1982;3:405-9.
2. Klevens RM, Edwards JR, Richards CL, Horan T, Gaynes R, Pollock D. Estimating health care-associated infections and deaths in US hospitals, 2002. Public Health Rep 2007;122:160-6.
3. Boyce JM, Pittet D. Guideline for hand hygiene in health-care settings: recommendations of the Healthcare Infection Control Practices Advisory Committee and the HICPAC/SHEA/APIC/IDSA Hand Hygiene Task Force. Infect Control Hosp Epidemiol 2002;23(12 Suppl):S3-40.
4. Pittet D, Hugonnet S, Harbarth S, Mourouga P, Sauvan V, Touveneau S. Effectiveness of a hospital-wide programme to improve compliance with hand hygiene. Lancet 2000;356:1307-12.
5. Gilbert K, Stafford C, Crosby K, Fleming E, Gaynes R. Does hand hygiene compliance among health care workers change when patients are in contact precaution rooms in ICUs? Am J Infect Control 2010;38:515-7.
6. Scott, RD. The direct medical costs of healthcare-associated infections in U.S. hospitals and the benefits of prevention. http://www.cdc.gov/ncidod/dhqp/pdf/Scott_CostPaper.pdf. Accessed February 14, 2014.
7. Centers for Medicare and Medicaid Services. Hospital-acquired conditions (present on admission indicator). Available from: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HospitalAcqCond/index.html. Accessed May 12, 2014.
8. Kurtzman ET, Buerhaus PI. New Medicare payment rules: danger or opportunity for nursing? Am J Nurs 2008;108:30-5.
9. Grant AM, Hofmann DA. It's not all about me: motivating hand hygiene among health care professionals by focusing on patients. Psychol Sci 2011;22:1494-9.
10. Gould DJ, Chudleigh JH, Moralejo D, Drey N. Interventions to improve hand hygiene compliance in patient care. Cochrane Database Syst Rev 2007;(2):CD005186.
11. Spano R. Potential sources of observer bias in police observational data. Soc Sci Res 2005;34:591-617.
12. Rogers B, Goodno L. Evaluation of interventions to prevent needlestick injuries in healthcare occupations. Am J Prev Med 2000;18(4S):90-8.
13. Sass D, Twohig M, Davies W. Toward the dissemination and acceptance of behavioral interventions: reporting sample characteristics. Behav Ther 2004;27:101-2.
14. Billington E, DiTommaso NM. Demonstrations and applications of the matching law in education. J Behav Educ 2003;12:91-104.
15. Alling K, Poling A. The effects of differing response-force requirements on fixed-ratio responding of rats. J Exp Anal Behav 1995;63:331-46.
16. Bailey JS, Burch MR. Research methods in applied behavior analysis. Thousand Oaks, CA: Sage; 2002.
17. Tromp M, Huis A, de Guchteneire I, van der Meer J, van Achterberg T, Hulscher M, et al. The short-term and long-term effectiveness of a multidisciplinary hand hygiene improvement program. Am J Infect Control 2012;40:732-6.
18. Kazdin AE. Behavior Modification in Applied Settings. Pacific Grove [CA]: Brooks/Cole; 1994.
19. Morgan DL, Morgan RK. Single-participant research design: bringing science to managed care. Am Psychol 2001;56:119-27.
20. Sax H, Allegranzi B, Uçkay I, Larson E, Boyce J, Pittet D. 'My five moments for hand hygiene': a user-centred design approach to understand, train, monitor and report hand hygiene. J Hosp Infect 2007;67:9-21.
21. Huitema BE. The Analysis of Covariance and Alternatives: statistical methods for experiments, quasi-experiments, and single-case studies. Hoboken, NJ: John Wiley & Sons, Inc; 2011.
22. Hettmansperger TP, McKean JW. Robust Nonparametric Statistical Methods. A simple and powerful test for autocorrelated errors in OLS intervention models. Psychol Rep 2000;79:300.
23. Baum WM. On two types of deviation from the matching law: bias and undermatching. J Exp Anal Behav 1974;22:231-42.