

# Beyond a Good Story: From Hawthorne Effect to Reactivity in Health Professions Education Research

---

## **AUTHORS**

Elise Paradis, PhD.\* Assistant Professor, Leslie Dan Faculty of Pharmacy and Department of Anesthesia, Faculty of Medicine, University of Toronto; Scientist, The Wilson Centre, University of Toronto and University Health Network. 144 College St, Toronto, ON M5S 3M2 CANADA. [elise.paradis@utoronto.ca](mailto:elise.paradis@utoronto.ca)

Gary Sutkin, MD. Associate Professor, Department of Obstetrics, Gynecology, and Reproductive Sciences, University of Pittsburgh Medical Centre, Magee-Womens Hospital, 300 Halket Street, Pittsburgh, PA 15213 USA. [gsutkin@magee.edu](mailto:gsutkin@magee.edu)

\*Corresponding author.

## **INSTITUTIONS WHERE THE WORK WAS PERFORMED**

Leslie Dan Faculty of Pharmacy, University of Toronto;  
Department of Obstetrics, Gynecology, and Reproductive Sciences, University of Pittsburgh Medical Centre.

## **FINANCIAL SUPPORT**

Elise Paradis' data collection was supported by a grant by the Gordon and Betty Moore Foundation (2012-2013).

## **KEYWORDS**

Qualitative research methods; clinical education; qualities/skills/values/attitudes

## ABSTRACT

### *Context*

Observational research is increasingly being used in Health Professions Education (HPE) research, yet it is often criticized for being prone to the Hawthorne Effect, defined as a research participant's behavioral alteration in response to being observed. This article explores this concern.

### *Methods*

First, this article briefly reviews the initial Hawthorne studies and the original formulation of the Hawthorne Effect, before turning to contemporary studies of the Hawthorne Effect in HPE and beyond. Second, using data from two observational studies (in the OR and in the ICU), this article investigates the Hawthorne Effect in HPE.

### *Findings*

Evidence of a Hawthorne Effect is scant, and amounts to little more than a good story. This is surprising given the foundational nature of the Hawthorne Studies in the social sciences and the prevalence of our concern with observer effects in HPE research. Moreover, the multiple and inconsistent uses of the Hawthorne Effect have left researchers without a coherent and helpful understanding of research participants' responses to observation. The authors' HPE research illustrates the complexity of observer effects in HPE, suggests that significant behavioral alteration is unlikely in many research contexts, and shows how sustained contact with participants over time improves the quality of data collection.

### *Conclusion*

This article thus concludes with three recommendations: that researchers, editors and reviewers in the HPE community use the phrase "participant reactivity" when considering the participant/observer/research question triad; that researchers invest in interpersonal relationships in their study site to mitigate the effects of behavioral alteration; and that researchers use theory to make sense of participants' behavioral alteration and use it as a window into the social world. The term "participant reactivity" better reflects current scientific understandings of the research process and highlights the cognitive work required of participants to alter their behavior when observed.

*Perhaps the most important lesson to be learned from the original Hawthorne experiments is the power of a good story* (Levitt and List, 2011: p. 237).

## **Introduction**

Qualitative research has grown in both importance and quality over the past decades in the field of health professions education (HPE). Observational methods in particular are increasingly being used, yet a common critique persists and may prevent their full appreciation: that they are prone to the “Hawthorne Effect,” defined as research participants’ alteration of behavior when observed. HPE researchers should be concerned about observer effects on behavior, for if they exist and they are strong, then observational research – whether experimental or ethnographic – merely reports what research participants *want us* to see, rather than exposes *what really goes on* when no one is watching.

In this article, we argue that this is an over-simplistic understanding of observer effects in research. We first review the Hawthorne Studies, their critiques, and the original formulation of the Hawthorne Effect. Second, we provide an account of contemporary research on the Hawthorne Effect, efforts to make this methodological artifact visible, and recent examples of its application in the HPE literature. Third, we use examples from our own observational research to illustrate the complexity of observer effects on participant behavior in two settings: the operating theater and the intensive care unit (ICU).

In light of this evidence, we suggest that it is time for the HPE community to reconsider how observers’ presence impacts research findings, and suggest some fertile avenues for future research that consider research participants’ reactivity to the goals of research. We hope this paper will help us as a community move beyond perpetuating the good but inflated story of the Hawthorne Studies and engage in higher-quality observational research.

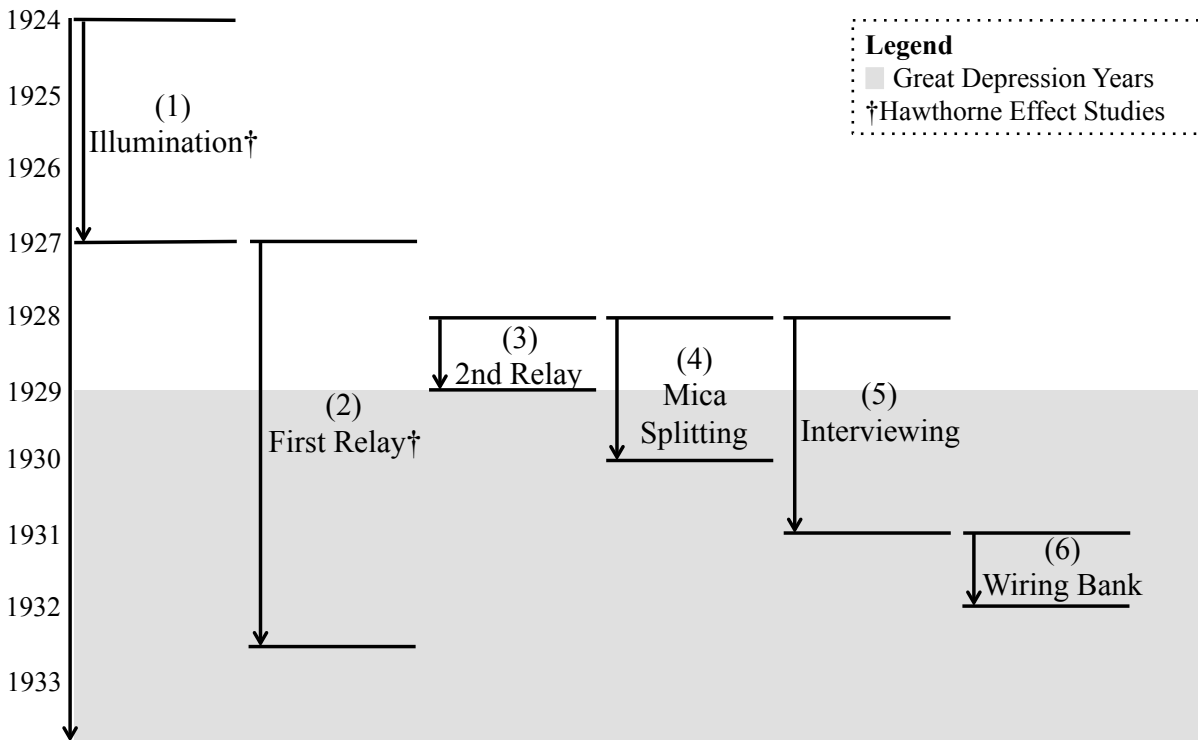
## **The Hawthorne Studies and the Hawthorne Effect**

The six Hawthorne Studies (see Figure 1) were conducted at the Hawthorne Works factory of the Western Electric Company in Cicero, Illinois, between 1924 and 1933.<sup>1,2</sup> The research team was composed of local researchers working in collaboration with Harvard Business School researchers including Elton Mayo, a pioneer of industrial psychology, and L.J. Henderson, a physiologist interested in industrial fatigue.<sup>1</sup>

The Hawthorne experiments were originally designed to study the effect of working conditions on productivity. They have taken on a “mythical power” in the social and behavioral sciences mostly because, according to Gillespie (1991), of the overarching “narrative of scientific defeat transformed into victory by a leap of insight.”<sup>1</sup> The Hawthorne story is generally told as follows. Scientists, after failing to establish a causal link between working conditions and productivity, finally “discovered” that “special attention by supervisors led to improved productivity.”<sup>3</sup> It is this special attention that is generally thought to constitute the Hawthorne Effect. After combing through the original Hawthorne archives, however, Gillespie has convincingly argued that this story misrepresents what actually happened:<sup>1</sup> Mayo was interested all along in documenting how “human factors” – over and above working conditions – were key to productivity. When local and university researchers failed to establish a causal link between productivity and one of the dozens of experimental changes they had introduced, they basically turned to their own, initial presuppositions about social behavior. There was no “ah-ha!” moment in the archives. Rather, the good story was constructed by the researchers in published materials.

----INSERT FIGURE 1 ABOUT HERE----

Figure 1: The Hawthorne Studies, 1924-1933



Source: Adapted from Franke and Kaul 1978. “The Hawthorne Experiments: First Statistical Interpretation.” *American Sociological Review*. 43(5): 623-643.

The Hawthorne studies have been the object of a wide critical scholarship from the 1950s onwards, and thus we can’t review it fully here. Instead, we provide a critical account of the two studies that purportedly demonstrated a Hawthorne Effect, the Illumination Study and the First Relay Assembly Test Room Study (see Figure 1).

According to academic folklore, the Illumination Study showed that every change in illumination of the factory floor increased productivity, even when it was reduced to the level of moonlight. Levitt and List (2011) summarized their reanalysis of the long-lost Illumination Study data by writing that “an honest appraisal of this experiment reveals that the experimental design was not strong, the manner in which the studies were carried out was lacking, and the results were mixed at best,”<sup>4</sup> and that “[p]erhaps the most important lesson to be learned ... is the power of a good story.”

The second and arguably most famous study was the First Relay Assembly Test Room Study, which investigated the productivity of five women operators working under close research supervision. It ran for 24 experimental periods and was terminated mid 1932 because of

the Great Depression. The story usually told about this study is that regardless of what the experimenters did, productivity increased: there were no connections between productivity and length or frequency of breaks, quality of the lighting, length of the work day, physiological markers, etc. The original<sup>5</sup> and secondary<sup>6</sup> published accounts typically hide the crucial fact that two of the initial operators were forcibly removed for insubordination and replaced by much faster and much more motivated operators. The net and immediate result was a massive growth in group output that continued throughout the study.<sup>7</sup>

Two reanalyses of the First Relay Assembly Study further undercut the traditional account. Franke and Kaul (1978) found that three variables explained about 90% of the variance in the productivity of the operators: the aforementioned replacement of two operators for more productive ones, the economic adversity introduced by the Great Depression, and the quality of the raw materials provided to operators.<sup>7</sup> Similarly, the reanalysis of the original data by Jones (1992) confirmed the absence of a Hawthorne Effect.<sup>8</sup>

--- INSERT BOX 1 ABOUT HERE ---

**Box 1: French's Canonical Description of the Hawthorne Effect**

“From a methodological point of view, the most interesting finding was what we might call the “Hawthorne Effect.” In order to manipulate more precisely the physical factors affecting production, the experimenters had set up a special experimental room for a small group of girls (sic!) who were wiring relays. This wiring room was separated from the rest of the factory, and the girls working in it received special attention from both outside experimenters and the management plant. Careful studies of this wiring group showed marked increases in production which were related only to the special social position and social treatment they received. Thus, it was the “artificial” social aspects of the experimental conditions set up for measurement which produced the increases in group productivity.”<sup>9</sup>

We will conclude this section by reconsidering the origins and initial meaning of the phrase “Hawthorne Effect.” The original accounts made no mention of the Hawthorne Effect, and the phrase has been attributed to John R.P. French, author of a chapter on field experiments a key research methods textbook of the era.<sup>10</sup> We reproduced what can be said to be the canonical definition of the Hawthorne Effect in Box 1. Most authors since French have truncated the excerpt and used only this sentence to define the Effect: “Careful studies of this wiring group showed marked increases in production which were related only to the *special social position and social treatment* [the women] received” (our emphasis). Taken out of context, this quote doesn't do justice to the argument put forth by French in the rest of his chapter, and cannot account for what French might have meant by “the ‘artificial’ social aspects of the experimental conditions.”

A closer reading of the chapter helps frame what French might have meant by the Hawthorne Effect and the artificial social aspects of experimental conditions. Indeed, French's chapter contrasts laboratory experiments with field experiments, which he defines as “a

theoretically oriented research project in which the experimenter manipulates an independent variable in some real social setting in order to test some hypothesis.”<sup>9</sup> According to French, a main limitation of findings obtained in the laboratory depends on the fact that “social behavior is role-determined.” Since participants in laboratory experiments are generally preoccupied with “being a [research] subject,” findings from the laboratory “cannot be generalized to apply to other roles without additional research.”<sup>9</sup> In contrast, field experiments are held in “real” social settings, and thus more likely to yield transferable results. In this light, the excerpt in Box 1 emerges a critique of the “artificial” conditions created by the Hawthorne experimenters, conditions “so different from the usual factory conditions that this experiment stands midway between a field experiment and a laboratory experiment.”<sup>9</sup>

In summary, French’s Hawthorne Effect does not articulate a threat for naturalistic, observational studies of the sort typically conducted by ethnographers. Rather, it highlights the limitations of laboratory findings and their lack of generalizability. The decontextualization of the above quote thus appears partly to blame for our current understanding of observer effects.

### **Contemporary Studies on the Hawthorne Effect**

Given the mythical status of the Hawthorne Effect, one would assume that contemporary researchers must have proved its existence, overshadowing its absence from the original studies. Our review of contemporary studies suggests that evidence of a Hawthorne Effect is limited, however. Several contemporary authors have sought to identify the mechanisms underlying the Hawthorne Effect, in the hopes that they might prevent it or establish controls. Adair (1984) reviewed 40 educational and psychological studies claiming to control for a Hawthorne Effect and found that authors generally posited three mechanisms: special treatment or attention accorded subjects, awareness of being in an experiment, and change in routine or novelty. The reviewed studies reported no influence of these variables on behavior, however, which led Adair to conclude that the Hawthorne Effect “anecdotally fits situations we recall, yet on closer examination it is meaningless.”<sup>3</sup> To mediate and assess behavioral change, he suggested instead that researchers focus on how participants understand the research, define the task, and adapt their behavior accordingly.

----INSERT BOX 2 ABOUT HERE ----

<p><b>Box 2: The Case Against the Hawthorne Effect</b></p> <ul style="list-style-type: none"><li>- Data from the original Hawthorne studies are not compatible with a Hawthorne Effect.<sup>4,7,8</sup></li><li>- The original conception of the Hawthorne Effect<sup>9</sup> is generally discussed out of context; a contextual reading suggests broader methodological concerns about generalizability of experimental findings.</li><li>- The impact on behavior of the three main mechanisms identified by researchers as constitutive of the Hawthorne effect (special attention; awareness of being in an experiment; novelty) is not supported by research.<sup>3,11,12</sup></li><li>- Contemporary uses are so broad and inconsistent as to make the phrase meaningless.<sup>11</sup></li></ul>
---

A more recent review by social psychologists Chiesa and Hobbs (2008) has shown that authors continue to use the Hawthorne Effect to “refer to a wide (and often contradictory) range of phenomena.”<sup>11</sup> In the 227 introductory social science textbooks, encyclopedias and dictionaries they reviewed, authors used the phrase to bring attention to

1. contextual factors impacting the research process;
2. intervening variables in the observer/behavior alteration chain, including (a) “psychological” or “subjective factors” as well as (b) more “social” intervening variables such as improved interactions and changes in the social environment;
3. resulting changes in behavior among study subjects.<sup>11</sup>

Such broad and inconsistent use of the Hawthorne Effect led them to suggest that we should stop using the phrase, for it obfuscates rather than clarifies research.

Finally, a 2014 systematic review by McCambridge et al. used a broad definition of the Hawthorne Effect as one that “concerns research participation, the consequent awareness of being studied, and possible impact on behavior”<sup>12</sup> to evaluate whether it exists, under what conditions it operates, and the size of the effect if it exists. Out of 1,130 screened studies, only 19 met inclusion criteria, and 17 of these came from healthcare research. The authors found that randomized controlled trials provided evidence of small, statistically significant effects, but no proof of a consistent or coherent effect; that quasiexperimental studies were inconclusive; and that observational studies suggested that “the size of any effects of health-care practitioners being observed or being aware of being studied probably very much depends on what exactly they are doing, ... undermin[ing] further the idea that there is a single effect, which can be called the Hawthorne Effect.”<sup>12</sup> They suggest that one of the core mechanisms behind the Hawthorne Effect and behavior change might be “conformity to perceived norms or researcher expectations,”<sup>12</sup> echoing Adair’s earlier conjecture that the “context of an experiment provides cognitively capable adult subjects with the necessity to develop a purpose or meaning for the situation that they then use to guide their behavior.”<sup>3</sup>

Integrating the literature covered so far (see Box 2), we find very little support for a Hawthorne or Observer Effect except in situations where participants know about the research aims and adapt their behavior to meet what they perceive to be the observer’s expectations. Moreover, the confusion over what it actually refers to seems to hinder rather than help understand the limitations of observational research.

### ***The Hawthorne Effect in HPE Research***

Several scholars in the field of HPE have discussed the Hawthorne Effect. Ng, Lingard and Kennedy (2014) define the Hawthorne (or observer) effect as “when observed participants act differently from how they would act if the observer were not present.”<sup>13</sup> They identify two mechanisms whereby it can be minimized: 1) prolonged engagement in the field, time taken to establish trust or rapport, and observer comportment (including dress); 2) the establishment of a process for recording and reflecting on the Hawthorne effect during observations. In another research methods text, Boet et al. (2012) describe the Hawthorne Effect as a methodological concern to ponder alongside several others, and define it as “positively changed behaviour or performance resulting from awareness of being a part of a study.”<sup>14</sup>

Authors have tried to mediate the Hawthorne Effect using three main strategies. They have (1) placed multiple observers strategically in different operating rooms,<sup>15-17</sup> (2) collected data over prolonged periods of time and during long surgeries to encourage subjects to act naturally,<sup>15,17-20</sup> and (3) encouraged observers to record examples of the Hawthorne Effect in

their field notes.<sup>17</sup> The many occurrences of over conflict between participants reported in the above-cited research suggest that the observers' presence did not importantly alter behavior, and thus that participants are either not conscious of observer effects on their own behaviors, or did not report them to researchers.

Strangely, we were unable to locate examples of observer effects in published HPE articles, despite how it is an accepted methodological challenge for observational research. This absence led us to consider observer effects in our research, which is what we turn to next.

### **Considering the Hawthorne Effect in Our Research**

In this section we use some of our own observations to question the effect of observers on participants' behavior in the operating theater and in the ICU.

#### ***Observational Research in Surgery***

In the collaborative surgical education research that I (GS) have led<sup>21,22</sup> we typically place a videographer in the operating room and give her a handheld video camera. The observer is thus easily identifiable because the camera must be placed close to the surgeons to capture voice and subtle movements. We routinely interview the surgeons following the case and ask not only about their teaching, but also the question "were you aware of the videographer or the camera during the surgery?" Without exception, everyone has said they were not aware. By contrast, during our whole project we encountered only one example in which the surgical team began a discussion with the videographer about her presence:

#### **Vignette 1: If I had an invisibility cloak**

Attending and Fellow are operating together intently. The observer is directly across from the attending and has the camera focused on her. The attending looks directly at the observer and into the camera, pauses and asks "So is this your research project?"

Observer: "Yes."

Resident: "Are you allowed to talk to us during this?"

Attending: "Are you *not* allowed to talk to us?"

Observer: "Well, I prefer not to. But you can address me, that's okay."

[Laughs.] "I'm just trying to be as inconspicuous as possible."

Attending: "I can tell you you're not inconspicuous."

More laughter from all participants.

Observer: "If I had an invisibility cloak, I'd be better off."

Attending: "That's right you need an invisibility cloak."

The Attending immediately refocuses on the operative field. Only twice over 115 minutes of filming did any member of the surgical team look at the camera, despite the observer standing within 3 feet of the surgical field.

Despite generally noting that the surgical team ignores the researcher, we have been concerned that the presence of an observer recording the surgery would decrease the incidence of use of humor, especially off-color humor, and more animated, pointed teaching, yet we saw numerous instances of both, including an inappropriate comment directed from an attending physician to a medical student about his race.

We were also concerned that attending physicians would hide their typical hierarchical, demanding behaviors. Yet in one example, during an audiorecorded postoperative interview, the attending reflected on her behavior in front of the camera:



### **Vignette 2: I'm going to kill you**

There's one point where I say "If you goof up, I'm going to kill you." (I say that all the time. They know it's a joke. If there's a very serious step, I'll stop and tell them how to go through the motions. ... I don't think anyone really worries I'm going to kill them. (It communicates:) "This is an important step. This is something really, really important to remember.")

It is not clear whether she was aware that she made the threatening comment during the surgery, or whether it struck her later that this language was captured and that it thus needed to be defended.

In summary, we believe that observers are unlikely to have a significant impact on surgical teaching. The operating room is a high-tech, fast-paced environment filled with many participants and distractions. Surgeons are intently focused on the operative field, and the ease with which they block out distractors and stimulants – such as the nurses chatting or the music playing – suggests that they are not likely to be influenced by another quiet presence. The operating room is sometimes referred to as the surgical "theater," and anyone who has observed surgery knows that it is a performance and has an audience. Surgeons are thus used to surveillance: their activities are public and actively watched by many, including students, other healthcare professionals, sales representatives, and other observers. In such an engrossing context, we believe that observer effects might be dampened.

### ***Ethnographic Research in the ICU***

Here I (EP) draw on data collected during a year-long comparative ethnographic study of team interactions in four American intensive care units (ICUs).<sup>23,24</sup> There are many ethnographic approaches, but ours is best described as critical ethnography, which aims to challenge the status quo by describing how local culture and practices interact and intersect with broad social dynamics and individual histories.

Entry into a new social environment is difficult for all newcomers, ethnographers or not, and my first visits to the study ICUs were difficult as well. The initial work of introducing yourself, the study and its goals, is always taxing, as there is much to learn: people's names and what they do; how the division of labor is structured; formal and informal hierarchies; the routines, tasks, taboos, running gags, etc. that make a social space special. Once social connections with participants are made, however, the ethnographer has access to a fantastic source of "emic" data: data originating from insiders in the study site that can reflect how they see the world and why they act as they do within it. This emic perspective can then be combined with the ethnographer's "etic" perspective, which reflects more closely their own social origins, training, values and beliefs. Clearly, one of the key strengths of ethnography arises from this "deep embeddedness": knowing the clinicians at my field sites enabled me to witness situations that should arguably be hidden if the Hawthorne Effect was really strong, as in the following vignette:

### **Vignette 3: A feisty little lady**

Lynn (RN), Emily (RN) and Jessie (Certified Nursing Assistant) are working together to re-install Ms. Lee (in her 70s, delirious, and wearing a C-spine collar) onto her bed. Her twisting and turning landed her at the foot of her ICU bed. Joking, Lynn says: "Time not to torture old people!" They try to pin Ms. Lee down to lift her up on the bed, but she resists, and Lynn asks for restraints to help them do their work. Emily motions "They're over there" with her head before

lunging towards a chest of drawers. Ms. Lee kicks as the nurses struggle to put on the restraints.

Lynn: “You know Ms. Lee, it’s good to be strong.”

Emily: “She’s a feisty little one.”

Lynn jokes that Ms. Lee must have had “two glasses of plum wine at the restaurant,” leading to her fall and hospitalization. She then notices that Ms. Lee’s gown is lifted up to her belly, and voices what she perceives to be her patient’s emotions: “She would be very upset that she was displaying all her old glory” like that. She pulls Ms. Lee’s gown down and squeezes the blanket tightly under the mattress on both sides. “I don’t think it’s illegal” to restrain her movement like that, she ventures, before concluding: “Torture was approved!”

Lynn, Emily and Jessie were undoubtedly aware of the social norms that they broke during my observations, including a respect for all patients (and especially elders), and the adherence to hospital protocols and state legislation. Yet time and social ties seem to be important factors in dampening research participants’ response to the observer’s presence for three main reasons. First, we are more likely to act according to shared cultural norms in front of strangers than in front of people we know.<sup>25</sup> Second, the goal of the study is typically explained upon first encounter, and might fade from participants’ memory over time. Third, trust builds over time, as the ethnographic outsider becomes more of an insider. This special relationship probably has deeper implications for the ethics of ethnographic research than for the validity of its findings. What should the observer report to authorities, or even publish? What kinds of relationships are appropriate and legitimate between ethnographer and participants? Where does power fit in?

Let us now turn to another Vignette that illustrates the complexity of observer effects and the potential of ethnography.

#### **Vignette 4: Curtain’s up!**

Eric (an ICU Fellow) and Lauren (an ICU Resident) are working together to put a trialysis port on Mr. White, a homeless man with a beard. As Lauren goes over the case and the procedure’s different steps, Eric repeatedly corrects her. He then admits that he is “totally paranoid” about cleaning skin surfaces: “People are gross,” he says, “especially homeless people,” who have feces all over. His solution: a round of Chlorhexidine, one of soap and water, and two more of Chlorhexidine.

Once Mr. White is clean, Eric, who wants to leave by 6pm, says: “You’ve got three minutes left, so you should be good.” He then looks in my direction, and adds: “I was going to say something really sarcastic, but I know you’re listening.” I tell him to ignore me.

Lauren tells Eric that another ICU fellow teaches her differently: “*He* says: ‘it’s not going to be perfect, move on,’ while *you* say: ‘it is perfect, move on.’ It’s confusing.” Eric looks up at me: “I get written up [for bad behavior] all the time.” I laugh, and he reiterates: “I’m not joking.”

Lauren is looking at the ultrasound machine. When she finds the artery on the monitor, Eric cheers: “You found it right away, good job!” He then turns towards me: “Did you note that? Positive reinforcement. We’re going to be done in two minutes.”

Lauren struggles a bit, poking at Mr. White’s skin. Blood squirts as she pushes the wire in, and the patient makes a big jerking motion. An alarm goes off, and

she apologizes: “It was an idiotic move.” Eric (to Lauren): “It was a learner move, not idiotic.” (To me): “Did you get this?”

As I walk out of Room 6, Pamela (RN) tells Russell (Respiratory Therapist) that she sent me to watch Eric and Lauren. “Does it happen often?” I ask. “That blatant attitude? No,” responds Pamela, “it is just confirming that he is a capital D.” Russell explodes laughing and concurs.

This vignette once again shows the importance of relationships. Pamela, a trusted informant, encouraged me to observe Eric and Lauren, which led to the above observations. Interpreting Eric’s behavior in this vignette is quite complex, however, and sheds new light on observer effects. Eric knows fully that he is observed, and partly adapts his behavior by instructing me to write down the “positive reinforcement” he gave Lauren and flagging moments where he self-censored. In doing this, he shows his interpretation of meaningful clinical teaching, and of appropriate behavior. Yet Eric also admits that he gets scolded for bad behavior “all the time,” and his derogatory beliefs and impatience are visible on many occasions, as he calls homeless people “gross,” sets unrealistic expectations with Lauren, and is unable to absorb the feedback he receives on the incongruence of his instructions with those of another fellow.

One could imagine a scenario where Eric convinces me of his clinical teaching skills and exemplary behavior through skillful deception. If he had not, for example, ruined his performance of “positive reinforcement” by both flagging it to me and pressuring Lauren for time, I might have seen him as a good teacher. In such a scenario, my initial (etic) assessment would have been challenged by the emic information given by Pamela and Russell, who clearly dislike him. Ethnographic relationships thus partially work as guarantors against deception, as skilled ethnographers are able to refine their emerging understandings of situations with others in the field.

## Discussion

In this paper, we have shown that data from the original Hawthorne Studies are not compatible with a Hawthorne Effect, that the original framing of the Hawthorne Effect has been distorted by decontextualization, and that despite the very limited evidence of a Hawthorne Effect, contemporary researchers use the phrase indiscriminately. This is a damning portrait, and we believe that we have to discard the Hawthorne Effect as a useful scientific construct. The Hawthorne Studies told a good story, but it was not supported by data. We cannot, as a community, uncritically repeat “academic folklore”<sup>26</sup> that distorts our understanding of iconic studies and impedes science.

We suggest referring to research participants’ behavior modification as “participant reactivity” rather than the Hawthorne Effect. We define participant reactivity as a form of researcher/participant effect that comes from participants’ active engagement with the research and its aims, a process that leads to behavioral adaptation in line with perceived social norms. We name this effect participant reactivity in recognition that the main mechanism for behavioral change lies in participants’ cognitive work, not merely by the presence of researchers. Such an effect should be small, however, since social psychological research shows that “one’s actions are relatively unaffected by normative information—even one’s own—unless the information is highlighted prominently in consciousness,”<sup>27</sup> and that this type of information loses its salience over time.<sup>27</sup> Consequently, some contexts and tasks are more likely to be influenced by the presence of an observer, including social evaluative tasks<sup>28</sup> that include a clear right answer (e.g. hand washing<sup>29,30</sup>), especially if they are bounded in time.

While several HPE scholars have grappled with the Hawthorne Effect over the years, we were unable to locate examples of observer effects in this literature. We have thus tried to enrich our methodological understanding using data from our own observational research. On this basis, we argue that rigorous, reflexive research in HPE is unlikely to be distorted by strong observer effects, for three main reasons. First, the demands of care are often too great for providers to access normative information. When the work is routine, automatic responses are likely; when it is not, focus is likely to be on the task rather than on normative information. Second, health professionals are always observed. In teaching hospitals in particular, senior healthcare providers are consistently watched by trainees and by other healthcare professionals on their teams. Meanwhile, trainees are monitored regularly by peers, superiors, and other healthcare professionals. How and why the researcher's gaze would lead to further transformations of behavior is unclear, especially if the research question does not dictate the appropriate behavior. Third, researchers working within ethnographic traditions come to know their participants over time. Personal connections provide researchers with access to a broad range of situations, including those that display less-than-ideal behavior, and with information that protects against deception and manipulation.

Instead of asking about the Hawthorne Effect, researchers who use observations and reviewers of such research should try to articulate their methodological approach and concerns more specifically. We have laid out some specific suggestions for researchers and reviewers in Box 3. Ethnographers in particular need to describe more fully their relationships to participants and be more reflexive about their practice, something that they haven't done in the past.<sup>27</sup>

-- INSERT TEXT BOX ABOUT HERE --

### **Box 3: Observational Research: Advice for Researchers and Reviewers**

#### **Researchers**

In order to minimize your participants' reactivity to being studied:

- Embed yourself within the environment and conduct sustained observations.
- Cross-check your own (etic) understanding of situations with the emic view of participants. Establishing rapport with your participants will help.
- Record and evaluate interactions between subjects and researchers.
- Consider limiting your participants' awareness of your specific research question.
- Pay particular attention to tasks where the "right" behavior is known and easily enacted (e.g. hand washing, feedback, etc.).

#### **Reviewers**

Consider whether the participants' behaviors could have been impacted by reactivity to research, particularly if they were aware of the research question and could have easily conformed to researchers' expectations.

Finally, using theory will help researchers who use observational methods make sense of behavioral adaptation. As social beings who respond to their environment and those around them, research participants will sometimes alter their behavior or at least *perform* what they believe is expected of them, as we have seen in the case of Eric. We know, for example, that students applying to medical school adapt their behavior to match known expectations,<sup>31,32</sup> and that trainees enact what they believe is the right way to doctor.<sup>33</sup> Here the theoretical insights generated by social psychology<sup>27</sup> (including Erving Goffman's work<sup>34-36</sup>) are a good place to start.

Instead of seeing behavioral adaptation as a failure of observational research, we should embrace it for what it can teach us about the social world. Humans are not atoms, and observer effects in social science research are not equivalent to Heisenberg's Uncertainty Principle. Health professions education researchers must recognize that research with human beings is not likely to deal in certainty, but rather that it is delicate and multifaceted. This fundamental, constructionist view of behavior is the first step in recognizing and honoring the complexity of the social world, and thus a critical step to take if we are to transform health professions education.

## References

1. Gillespie R. Manufacturing knowledge: a history of the Hawthorne experiments. New York: Cambridge University Press; 1991.
2. Hassard J. Rethinking the Hawthorne Studies: The Western Electric research in its social, political and historical context. *Human Relations*. 2012;65(11):1431-1461.
3. Adair J. The Hawthorne Effect: A Reconsideration of the Methodological Artifact. *Journal of Applied Psychology*. 1984;69(2):334-345.
4. Levitt SD, List JA. Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments. *American Economic Journal. Applied Economics*. 2011;3(1):224-238.
5. Roethlisberger FJ, Dickson WJ. *Management and the worker: an account of a research program conducted by the Western Electric Company, Hawthorne Works, Chicago*. Cambridge, MA: Harvard University Press; 1939.
6. Bramel D, Friend R. Hawthorne, the myth of the docile worker, and class bias in psychology. *American Psychologist*. 1981;36(8):867.
7. Franke RH, Kaul JD. The Hawthorne Experiments: First Statistical Interpretation. *American Sociological Review*. 1978;43(5):623-643.
8. Jones SR. Was there a Hawthorne effect? *American Journal of Sociology*. 1992:451-468.
9. French JRP. Experiments in field settings. In: Festinger L, Katz D, eds. *Research methods in the behavioral sciences*. New York, NY: Holt, Rinehart and Winston; 1953:98-135.
10. Festinger L, Katz D, eds. *Research methods in the behavioral sciences*. New York, NY: Holt, Rinehart and Winston; 1953.
11. Chiesa M, Hobbs S. Making sense of social research: How useful is the Hawthorne Effect? *European Journal of Social Psychology*. 2008;38(1):67-74.
12. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *Journal of clinical epidemiology*. 2014;67(3):267-277.
13. Ng S, Lingard L, Kennedy TJ. Qualitative research in medical education: Methodologies and methods. *Understanding Medical Education: Evidence, Theory and Practice*. 2013:371-384.
14. Boet S, Sharma S, Goldman J, Reeves S. Review article: medical education research: an overview of methods. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*. 2012;59(2):159-170.
15. Lingard L, Garwood S, Poenaru D. Tensions influencing operating room team function: does institutional context make a difference? *Medical education*. 2004;38(7):691-699.
16. Lingard L, Regehr G, Espin S, Whyte S. A theory-based instrument to evaluate team communication in the operating room: balancing measurement authenticity and reliability. *Quality and safety in Health Care*. 2006;15(6):422-426.
17. Lingard L, Reznick R, Espin S, Regehr G, DeVito I. Team communications in the operating room: talk patterns, sites of tension, and implications for novices. *Academic Medicine*. 2002;77(3):232-237.

18. Kennedy TJ, Lingard L, Baker GR, Kitchen L, Regehr G. Clinical oversight: conceptualizing the relationship between supervision and safety. *Journal of General Internal Medicine*. 2007;22(8):1080-1085.
19. Kuper A, Nedden NZ, Etchells E, Shadowitz S, Reeves S. Teaching and learning in morbidity and mortality rounds: an ethnographic study. *Medical education*. 2010;44(6):559-569.
20. Reeves S. Community - based interprofessional education for medical, nursing and dental students. *Health & social care in the community*. 2000;8(4):269-276.
21. Sutkin G, Littleton EB, Kanter SL. How Surgical Mentors Teach: A Classification of In Vivo Teaching Behaviors Part 1: Verbal Teaching Guidance. *Journal of surgical education*. 2015;72(2):243-250.
22. Sutkin G, Littleton EB, Kanter SL. How Surgical Mentors Teach: A Classification of In Vivo Teaching Behaviors Part 2: Physical Teaching Guidance. *Journal of surgical education*. 2015;72(2):251-257.
23. Paradis E, Reeves S, Leslie M, et al. Exploring the nature of interprofessional collaboration and family member involvement in an intensive care context. *J Interprof Care*. 2014;28(1):74-75.
24. Paradis E, Leslie M, Gropper MA. Interprofessional rhetoric and operational realities: an ethnographic study of rounds in four intensive care units. *Advances in Health Sciences Education*. 2015;Online ahead of print. DOI 10.1007/s10459-015-9662-5.
25. Ridgeway CL. *Framed by gender: How gender inequality persists in the modern world*. Oxford University Press; 2011.
26. Todd JT, Morris EK. Case histories in the great power of steady misinterpretation. *American Psychologist*. 1992;47(11):1441.
27. Cialdini RB, Goldstein NJ. Social influence: Compliance and conformity. *Annu. Rev. Psychol*. 2004;55:591-621.
28. Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*. 2003;88(5):879.
29. Eckmanns T, Bessert J, Behnke M, Gastmeier P, Rüden H. Compliance With Antiseptic Hand Rub Use in Intensive Care Units The Hawthorne Effect. *Infection Control*. 2006;27(09):931-934.
30. Maury E, Moussa N, Lakermi C, Barbut F, Offenstadt G. Compliance of health care workers to hand hygiene: awareness of being observed is important. *Intensive care medicine*. 2006;32(12):2088-2089.
31. Wright S. Medical school personal statements: a measure of motivation or proxy for cultural privilege? *Advances in Health Sciences Education*. 2015;20(3):627-643.
32. Broadhead RS. Individuation In Facework: Theoretical Implications From A Study Of Facework In Medical School Admissions\*. *Symbolic Interaction*. 1980;3(2):51-68.
33. Haas J, Shaffir W. Taking on the role of doctor: A dramaturgical analysis of professionalization. *Symbolic Interaction*. 1982;5(2):187-203.
34. Goffman E. *The presentation of self in everyday life*. Garden City, NY: Doubleday Anchor Books; 1959.
35. Goffman E. *Interaction ritual; essays in face-to-face behavior*. Chicago, IL: Aldine Pub. Co.; 1967.

36. Paradis E, Webster F, Kuper A. Medical education and its context in society. In: Walsh K, ed. *Oxford textbook of medical education*. Oxford, UK: Oxford University Press; 2012:136-147.