Introduction: The Housekeeping Work of IRB Meetings  
Who can you trust? In everyday life, people regularly distinguish between more and less trustworthy individuals. In medical contexts, patients and research participants build impressions of the trustworthiness of caregivers and clinicians, and these impressions shape their decisions. It is tempting to think that people in certain official roles do not—or at least should not—allow these kinds of impressions to affect their decisions. Researchers might expect that members of human-subjects review committees, for example, would not take informal, impression-based factors into account when evaluating investigators’ study protocols. After all, the tasks of board members are well outlined in the regulations that describe and govern the work of human-subjects committees (known as institutional review boards [IRBs] in the United States, research ethics committees [RECs] in the United Kingdom, and research ethics boards [REBs] in Canada). Trustworthiness is not an explicit criterion for the assessment of the risks of research for subjects.

Over the course of one year, I studied the full-board meetings of three IRBs and was surprised by the results: review-board members do evaluate investigators’ trustworthiness, and these evaluations have tangible consequences for investigators. Board members actively search for signs of investigators’ trustworthiness, and their informal judgments grow out of a seemingly mundane and benign processes in research review—what board members called “housekeeping work.” This article documents how housekeeping work creates advantages for some investigators while disadvantaging others.

Sociological Perspectives on Trust: IRBs as Auditors and Trustees

Trust is marked by an inclination to accept the appearance of situations and the honesty of people’s statements with no further evidence. Social scientists traditionally believed that levels of trust declined in modern societies as bureaucracies and mass culture grew. In recent decades, however, theorists have reconsidered this understanding of trust, arguing that interpersonal trust is not an exclusive feature of exchanges with individuals close at hand but is also a feature of networks of people.

In this vein, sociologist Anthony Giddens argued that modern bureaucracies, far from eroding trust, actually enable people to trust strangers and to overlook the uncomfortable reality that
on any given day the mundane activities of life could kill us. People rarely pause to wonder whether the strangers who prepare and package our food, or maintain our cars and airplanes, are doing their jobs. Government systems and professional organizations give people the sense that others are performing their due diligence and keeping the promises that service providers and bureaucrats make to strangers.

To create this sense, modern institutions build “audit cultures” that encourage people to develop techniques for holding one another accountable. An emblematic practice of audit cultures is the performance of spot-checks: health inspectors turn up at restaurants and records monitors drop by medical research offices. In theory, the possibility of an audit encourages self-surveillance. At the same time, people’s knowledge that their performance will be evaluated according to the standards of the auditors encourages people to match their behavior to auditors’ criteria without deeper reflection—to teach to the test, as it were. In many ways, IRBs exemplify the audit culture and its ironies. IRB members check up on investigators (or simply advertise the fact that they could spot-check them), thus reminding investigators that they are accountable to an organization for their actions.

To make sure that the standards of the outside auditors are met, organizations empower “trustees” to ensure that their employees are following the rules. Legal scholar Susan Shapiro explains that “many of these policing activities are found within trust organizations and institutionalized in the functions of compliance officer (brokerage firms), inspector general (government agencies), fact checker (magazines), internal-affairs division (policy departments), morbidity-mortality review committee (hospital), quality-assurance review (public accounting firms), and the like.” The specter of “outside” audits encourages organizations to recreate those audit practices within their own structures.

The fates of IRB members are entwined with the decisions of the people they are auditing, which makes IRBs the trustees of universities, hospitals, and other research organizations. They are groups within the organization that serve both to enact regulations and also to protect the organization from outside regulators by simulating, or pre-enacting, the work of those outside regulators. IRB members do not simply deflect accountability from themselves onto researchers; board members are implicated in researchers’ poor choices. Thus, IRB members become all the more responsible for investigators’ actions after they endorse the people and their practices. As board members imagine what the future will hold for themselves in tandem with an investigator, their biggest challenge as both auditors and trustees is to gauge the trustworthiness of the investigator.

Research Methods

This article is based on twenty interviews with IRB chairs at major research universities across the United States in 2003 and 2004, and on observations of meetings and interviews with members of three university IRBs between March 2004 and October 2005.

For interviews with IRB chairs, I drew a random sample of 20 percent of the 151 universities (n = 30) categorized as “doctoral/research universities—extensive” according to the Carnegie Classification of Higher Education, 2000 edition. I completed 20 interviews, a response rate of 67 percent. Interviews with this national sample were used to get a broad view of the common issues and modes of operation of IRBs in the United States.

I contacted three of the IRB chairs whom I interviewed as part of the
national sample to ask whether they would permit long-term observation of their full-board meetings. These three IRBs were chosen because their institutions were within a day’s drive and because their chairs had encouraged my research. Using this recruitment method raises the possibility of selection bias, but the direction of this bias is likely in favor of boards that are well-functioning and open to reflection about the way they made decisions. Thus, the sample over-represents self-aware and self-critical IRB members, the best-case boards from researchers’ perspective.

With IRB members’ consent, I attended the monthly full-board meetings of two IRBs for one year between 2004 and 2005 and audio-recorded most of the meetings. I also observed but did not audio-record the twice-monthly meetings of one IRB at a medical school for five months. The medical school board was one of several IRBs at a university that had additional boards for nonmedical research. This IRB, referred to here as the Adams Medical Board, met every other week for three to four hours, during which time I took field notes. I also interviewed 10 of the 11 regular members of the board, plus 1 nonvoting administrator. The other two IRBs I observed were the only boards at universities without medical schools (although investigators conducted vaccine trials, physiology studies, and other medical research often in cooperation with local clinics). In the case of the board, described here as Greenly IRB, members gave permission to audio-record their monthly meetings for one year. I supplemented these recordings with handwritten field notes and interviews with 11 of 14 board members. At the other university board, Sander State IRB, members granted permission to audio-record their monthly meetings after my fifth month with the board. Thus, I recorded meetings, which averaged just less than two hours, for the remainder of the year (seven months) and continued taking handwritten field notes. I also interviewed the 12 regular members of the board over the course of the year. In total, my observations of meetings of three IRBs were supplemented with recorded interviews with 34 of their members.

**Housekeeping Work: Judging Investigators’ Trustworthiness**

Conventionally, IRB members are thought to evaluate the content of study protocols. Through informed consent provisions, for example, board members judge whether researchers plan to tell participants that they will be studied and whether researchers intend to share all of the information that might be relevant to people as they decide whether to enroll. IRB members also consider the quality of investigators’ research design and the value of potential findings. In addition, IRB members—especially those in administrative roles—ask whether investigators are following the nuts and bolts of national regulations and local policies. Taken together, these practices map onto legal requirements for human-subjects protections and onto ethical imperatives such as autonomy, beneficence, and justice.\(^{16}\)

Yet IRB members look for more than content in documents: they also use documents to judge the investigator’s overall precision and carefulness. Board members do this because their approval of a given study is also a tacit statement that they—the legal trustees of the institution—believe that investigators will carry out studies in the way they describe. The IRB is not only evaluating investigators but also being evaluated by the federal government based on how their researchers conduct themselves going forward.

As a result, IRB members develop techniques to gauge the trustworthiness of
investigators. What many IRB members called “housekeeping work” involved using the documents that investigators submitted—such as consent forms and study protocols—to generate an impression of investigators’ carefulness and credibility, to use as a measure of the investigators’ trustworthiness in future interactions with research participants. When board members made housekeeping comments, they queried aspects of an investigator’s application that would not directly affect the research plan or the ability of potential subjects to decide whether to enroll in a study. They corrected typographical errors, discussed investigators’ formatting problems, and praised elegant-looking materials. In one meeting at Sander State, a physiologist serving on the board explained that he wanted to ask “a trivial question.” The word “lit” appeared in the application, and the physiologist asked, in light of the grammatical context, “What does this [lit] mean?” Other board members had the same question. The investigator clarified that the word should have read “list,” remarking, “It’s a typo!”

Housekeeping work was indispensable for IRB members, even when their edits had no effect on the study design or participants’ understanding of the research. Board members used the apparent degree of care taken in submitting a tidy application as a proxy for an investigator’s self-discipline and fastidiousness. For example, during one meeting of the Greenly IRB, a board member, Dr. M, criticized an application at length after pointing out that the researcher misspelled “principal” in the subject heading of his protocol. I had scheduled an interview with Dr. M after the meeting, and I asked how he went about reviewing protocols. Among other things, he looked for “misspellings” and other “editorial things” that “bother me” because such shortcomings demonstrated that the investigator had “a lack of attention to detail.”

Dr. M was looking at the mechanics of the documents, not simply their content. He read the documents for information about the study, but also for information about the quality of the researchers—two separate targets of evaluation. As he explained, an investigator’s housekeeping mistakes “colors the impression of the reviewer immediately” because it provides evidence that “this is a person who’s not careful enough to make sure that the word is spelled right.”

Sloppy proposals, in Dr. M’s view, were written by sloppy investigators, who had the potential to threaten the well-being of research participants. For Dr. M, “if [a researcher’s] attention to detail is not sufficient to know that in the major heading the words aren’t spelled right, I’m worried about [other things as well]. Do I have to read this thing carefully enough to make sure that all the doses, for example, are correct, that they’ve written the protocol correctly?” As a result, Dr. M and others like him would be primed to wonder, “Where else is it sloppy? [Does the researcher really mean] four micrograms of nitroglycerin instead of point-four or four hundred micrograms of nitroglycerin? How careful do I have to be?” By this logic, a typo in a clinical-study application might flag other errors with drug dosages and interactions.

Similarly, a historian on the Sander IRB described herself as “a stickler for detail” in protocol reviews. In addition to issues of confidentiality, she was particularly attuned to “any inconsistency in the protocols, any of the specifics.” She explained:

If it is an excessively sloppy proposal, I’m going to be more questioning about it. Even if the researcher thinks [the
study] is potentially valuable, I do think they should be made to take care, take the time, get it right. . . . I would be prejudiced against it if it is full of typos, inconsistencies, factual errors that would make me doubt. I’d be questioning about the ability of the researcher.

This interviewee indicated how board members could interpret a researcher’s apparent inattention to detail in documents as a measure of more general professional weaknesses.

For the IRBs, a researcher’s good or bad housekeeping of documents specifically suggested the amount of care and supervision she would direct toward the proposed study in the future. One IRB chair in the national sample explained how it is evident in an IRB application if an investigator has a general work style that involves impatience or carelessness. The board had received several seemingly rushed applications, which board members read as a “lack of oversight by some researchers.” Yet the IRBs’ efforts to clean up these applications were not only unrewarded, he felt, but actually counterproductive, because their housekeeping efforts gave investigators the sense that “we are a mindless bureaucracy that is just creating problems for them.” This was not the case, of course: IRB members felt they were doing the important work of both ensuring against consequential mistakes and developing an overarching sense of the reliability of the investigator in whom they were trusting.

Other board members said that the care with which documents were prepared was a reflection of the quality of an investigator’s overall study management. At Adams Medical, a clinician on the IRB explained, “The PI [primary investigator] has to sign off on everything in [the electronic submission system]. They have to approve it.” As a result, an untidy proposal suggested that the investigator either was (and therefore is) careless or was (and therefore is) a poor supervisor. Board members at Adams Medical expected the materials to be drafted by nurses and technicians in relatively low-wage positions who had little experience and needed to be trained. Nonetheless, he explained, “It’s always been very clear to me whether the PI actually looked [at the documents] or didn’t because sometimes the answers to the questions the IRB might write [to the investigator], or the way [investigator’s] responses are phrased, are completely off base. Or they’re not written in English, literally, or they’re flippant, or just evasive.” He continued, “You just see a lack of supervision, and that to me is a red flag. Then you get concerned that the patients are at risk. It’s not just administrative. It’s that they really don’t know what’s going on with the patients on those trials.” To IRB members, untidy documents suggested a dangerous future for research participants, the investigator, and the IRB itself as a trustee of the institution.

Although critics of IRBs often describe boards in opposition to investigators, IRB members saw an investigator’s actions as reflections of the board’s legal and ethical integrity. In practical terms this manifested itself in IRB members’ worries that federal regulators would cite or shut down research at their institutions because of an investigator’s protocol violation. Investigators’ paperwork took on a layer of importance because, as one board chair explained to me, unkempt documents might indicate “the PI trusts the research assistant to do things as per the protocol and doesn’t check up.” The IRB’s literal reading of the trustworthiness of an investigator in the PI’s paperwork could prevent “protocol violations and things like that [because the documents show that] the researcher
is getting too busy to pay attention to the details.” At Sander State, one board member caught a typo in the application that would have caused such a problem. The investigator had mistyped one digit of a number, which had been the age of inclusion for research participants. (He typed “28” instead of “18.”) Had the typo gone through, investigators could have been cited for protocol violations for all of the participants they enrolled between the ages of 18 and 28.

IRB members invest energy in judging investigators’ precision precisely because members cannot oversee all studies as they are carried out. In the process of review, IRB members are deciding whether to wager (perhaps lives and certainly the legal fees of their institution) that an investigator will act carefully in the future. When board members are skeptical, they intervene in a researcher’s work, for example, by requiring the investigator to report to the board earlier than legally required, or by asking investigators to file preliminary data with the board. Alternatively, board members might select the study for audit. An IRB chair at one top U.S. research institution explained:

We have a subgroup of the board that actually goes and physically audits what is done. So we do three or four a year. We’ll pick certain protocols that we want to have audited and they’ll go out and actually see how it was done, and whether it was carried out appropriately. It’s not that we don’t trust everybody. It’s just a good experience for all concerned.

Some would disagree with this chair’s view that auditing is good for all involved. Legal scholar Mark Hall argues that in the case of physician oversight, surveillance paradoxically creates mistrust even though the aim is to demonstrate compliance with rules and evidence of good behavior.  

In the preceding interview with the IRB chair, I followed up by asking how he selected the studies to be audited. “Well that’s an interesting question,” he replied. “The full board usually discusses protocols we’ve looked at that we think would be good ones to look at again, as far as: Is there anyone that was denied and now subsequently has gotten approval?” IRB members’ initial suspicions about a proposal ripple through the IRB relationship. “It’s not meant to be a punitive audit,” he assured me, “It’s just an audit to see how things are done.” It was, in other words, surveillance in action.

When IRB members read an application that required a good deal of housekeeping work, they often came to think of the investigator as having a flawed character (however minor) and not simply a flawed application. Subsequently, these investigators received closer scrutiny than others. As trustees of institutions, board members tried to prevent the consequences of investigators’ future carelessness and poor supervision from manifesting itself in the research itself.

Who Is at a Disadvantage because of Housekeeping Work?

In the United States today, fewer than 9 percent of IRBs invite investigators to attend meetings, and historically this number has been low. In 1978, a quarter of IRBs reported that investigators always attended meetings at which their proposals were discussed.  Twenty years later, approximately one-third of investigators attended meetings. This 1998 survey found that 42 percent of low-volume IRBs and 17 percent of high-volume IRBs “routinely encouraged [PIs] to attend the meetings or to be reachable by telephone.” Yet the invitations were rarely accepted. In low-volume boards, the study found, 22 percent of investigators “attended meetings or were
on call only when requested”; this was true of 41 percent of investigators in the case of high-volume boards. Historical data show that IRBs today are less likely than they were a few decades ago to invite investigators to the review of their protocols.

When investigators do attend meetings, they provide IRB members with a layer of information that is intangible and ephemeral. In the meetings I observed, board members sometimes solved problems quickly and with good will when investigators attended the meetings. For example, at one meeting in which the investigator was present, board members questioned the investigator about his application, which was riddled with typos:

Nigel (humanities professor): I have a linguistic curiosity here. On page two in the middle it says “AIM is a racial and surreal.”
IRB members: (Eighteen seconds of laughter and jokes.)
Nigel: Wrong expression? Is that what you meant?
Investigator: I meant “socio-economic.”
Owen: Spell check apparently liked it!
IRB members: (Five seconds of laughter.)

With the investigator present, IRB members laughed, commiserated, and, finally, trusted the investigator. The IRB approved the proposal, typos and all.

Yet face-to-face meetings did not always reassure board members of investigators’ trustworthiness. Often the type of investigator who appears trustworthy to IRB members resembles members of their social networks. The race, class, and gender composition of those who are selected to serve on IRBs (mostly white men as of 2002)20 seem to pattern IRBs’ evaluations of investigators.21

In the IRB meetings I observed, the presence of investigators did not always encourage more legitimate grounds of evaluation. For example, in one meeting a Korean woman attended the review of a study on which she was collaborating with two senior colleagues. She had attended the previous meeting, in which the board asked her to resubmit the application and consent documents. On her second visit, however, she still had not satisfied IRB members. “She revised it according to what we asked her to do before,” the board administrator remarked, but somehow “I don’t think she hit all the targets.” Others agreed.

The root of the problem, they felt, after seeing her in person, was that she did not understand English well enough. As one board member stated, “I don’t think she understands even what we’re saying. I really don’t.” For some IRB members, her status as a junior faculty member and a minority seemed inextricable from their legitimate concern about language skills and thus her ability to conduct the study. The community member tried to address the issue head-on with tact and “without denigrating her intelligence.” He likened his reading of the investigator’s documents to reading a bad translation of a classic novel. The crux of the problem, he explained to other members at the meeting, was that “her native tongue is Korean, which is a long, long way from English.” He used his role as a community member on the board (he was a white retired minister) to claim that because this investigator was a non-native English speaker, participants would not be safeguarded from research risks, even though two senior investigators were collaborating on the study: “What I’m saying is that I don’t think she appreciates the difference between what’s intrusive and what isn’t intrusive, simply because she’s operating outside of her frame of reference.”

The chair agreed with the community member’s suggestion that the IRB
should tell her collaborators about the problems they had working with her. He felt that “there’s no harm, and in fact, it may be positive to keep the other two investigators informed because they are listed as investigators, too.” Not everyone agreed. “I think that’s problematic,” the female administrator commented. (“Really?” the community member replied.) Another woman on the board, an exercise physiologist, explained, “It’s not like she’s their graduate student.” But the chair persisted, saying, revealingly, “But in a sense they’re being reviewed here.” In effect, the board members were making and dismantling investigators’ local reputations.

A 2008 study examined how the reviews of one IRB changed when investigators were invited to attend meetings. When investigators came to meetings, the IRB took fewer days to give a final decision about a protocol and required fewer meetings to come to an agreement. This study suggests that face-to-face interactions make reviews more efficient by speeding time to approval, even though in-person meetings early in the review process were more time consuming for investigators.

It is hard to disagree with greater efficiency, and it is likely that face-to-face meetings streamline housekeeping work by enabling investigators and board members to talk through typos. Without investigators present, IRB members must use proxies—most often documents—to gauge the trustworthiness of researchers. The efficiency of in-person meetings could be attributed to a reduced chance of mistaken impressions and to greater trust in investigators—but, as my evidence shows, this holds only for some kinds of investigators. Those who spoke more easily with the IRB tended to have the flaws in their application documents glossed over, although the errors might have been interpreted as a sign of a deeper weakness in the investigators had they not been socially and linguistically fluent. Investigators who did not share the dominant language of IRB members (English, in my cases) generated mistrust and suffered the consequences—potential audits and reputational damage among colleagues—because they both made grammatical mistakes and proof-reading errors in English that were interpreted as weaknesses for an investigator and then fell short in repairing their documents’ flaws in conversation with board members.

Conclusion: How Do We Trust?

My study of IRB meetings documents that trust is located and generated within specific social contexts. Building on this point, my empirical findings suggest that the puzzle of trust is best pieced together not by asking “who do we trust?” but by asking “how do we trust?” In other words, what tools and techniques do people use to determine trustworthiness within a given social setting? In the case of IRBs, housekeeping work is an important vehicle for assessing trustworthiness.

IRBs must assess protocols according to the criteria of 45 CFR 46, but the methods used to interpret these regulations emerge as IRB members work together to reach decisions. In order to understand and improve the way IRBs work to protect human subjects, it is important to consider how these secondary rules emerge and how they shape IRB decisions. The finding that housekeeping work advantages native-language speakers is worrisome, especially given the increase in language communities involved in transnational research protocols and in domestic studies.

The federal government of the United States is in the process of overhauling regulations governing research. The most effective reforms would shift the practices of deliberation and open the review process to new groups,
including language experts or members of new language communities.

As this article showed, housekeeping work is a technique for assessing trust that hides in plain sight. It may seem to be a silly, even embarrassing, part of a day’s responsibilities. Yet by using the paperwork that investigators complete and assemble, IRB members evaluate the people behind the paper. When investigators were not present in the flesh, the stewards of public trust used documents as tools to assess trustworthiness. IRB members judged the quality and character of investigators by the cleanness and correctness of the documents they submitted. Whether revealed in a text or regarded across a conference table, IRB members’ judgments rested on what they took to be evidence of investigators’ character and the future safety of research participants and their organization.

Notes