Article

Data Access as Regulation

Susan M. Sterett[1]

Abstract
This article considers calls for data transparency as research regulation and accountability. Rather than arguing for or against the value of sharing data, the article argues that understanding the call for data sharing requires questioning assumptions embedded in the debate about the context of scholarship and rethinking the purposes of data access. The article first argues that the spread of information available digitally means that researchers in the academy and outside it work with digital information, quite apart from mandates for data access. Second, replication as an accountability measure is often offered as one reason for making data available. However, scholars of replication have argued that replication has multiple components, many difficult to enact. Demands in universities for grant funding, impact by standard metrics, and newsworthy research encourage rapidly produced scholarship and research that makes big innovative claims. However, replication imposed sporadically cannot regularly counter these systematic incentives. If one purpose of data access is to regulate the research enterprise, scholarship on regulatory strategies and the difficulty of accomplishing goals via mandates illuminates the call for data access. Replication operates as a threat, one seen to generate incentives for good science, but is erratically enforced. Borrowing from the scholarship of audit and regulation, the article uses regulation, including audit, as accountability to argue that the sciences might need to address fundamental concerns about trust.

Keywords
data access, regulation, digitization, data ethics

Introduction

The rapid spread of digitization enables collecting, cleaning, archiving, sharing, and analyzing data at a distance from a field. Meg Leta Ambrose has named the process

[1]School of Public Policy, University of Maryland, Baltimore County, Baltimore, MD, USA

Corresponding Author:
Susan M. Sterett, School of Public Policy, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA.
Email: ssterett@umbc.edu

that turns information into uniform and readily shared points datafication (Ambrose, 2014: 10). Digitization has accompanied increasing calls for data access to promote research transparency, evidenced by guidelines in the United States from the National Science Foundation (NSF) and in the United Kingdom through the Economic and Social Research Council (ESRC) (n.d.), and boxes on journal submission sites to list where one could find that data used in an article. Calls for data sharing argue that sharing will serve multiple purposes, including facilitating replication and thereby incentivizing good scholarship, making material more broadly available for teaching, and checking on cheating or mistakes. In political science, economics, and social psychology, pressure for data access from funders and journals and people who analyze work to check on it have proliferated. Reasons for data access shift. Purposes include sharing to allow greater use of data in teaching, and allowing use by a broader public, including students, journalists, and citizens, as well as the familiar purpose of replication. Some anthropologists have argued that greater transparency about field sites would allow establishing baselines for later studies (A. Murphy & Jerolmack, 2016). The U.S. NSF has argued that if the public pays for data, then the public should have access. Statements from practitioners in different disciplines have also proliferated, from discussion about how to make research with sensitive data accountable to urging acceptance of diversity in practices (Freese & Peterson, 2017).

Rather than arguing for or against data access for any kind of work, this article considers data access as a form of regulation of the research enterprise. Regulating for research accountability implies a principal, to whom an agent answers. If the agent is the researcher, the principal is not very clear in many of the calls for data access. The agent could be the profession, or the funders who might pay for research, or those who might act on research, or the ideals of science. These could all be principals, and this article does not try to decide among them. The principal/agent problem presumes the agent has reasons not to do what the principal wants done. This problem has been widely used as a way to understand political accountability across institutions. Figuring out how to hold the agent accountable is difficult, particularly, when the agent has multiple principals. Employers, funders, and science are all principals. They work on different timelines rely on different incentives. Employers work with annual reports, and the employment market can reward dramatic findings. Funders work on funding cycles. Science can work over a century: The profession's rethinking of Margaret Mead's work (Shankman, 2009) and of the Stanford prison experiment (Kulig, Pratt, & Cullen, 2016) harken back to data collected decades ago. The Stanford prison experiment is now widely questioned, but the people who conducted the research ultimately built successful careers by common metrics: citations, and prestige of an employer. If the need to build a career is ever in tension with accountability to building knowledge, scholarship has a principal/agent problem.

Regulation and Work Practices as Ethical Concerns

Data sharing is a form of regulation that aims manage principal/agent problems. Therefore, the ethics of the context of work are integral to the ethics of data sharing.

Objections to data sharing come not only from concern for research subjects. Data access in the big data era is access in an era of digitization. That context reshapes relationships yet is often only a background in the data access discussion. Debates over the ethics of data sharing often center on obligations to research subjects, and they concern individual-level ethical decisions scholars make. However, researchers make those individual-level decisions in a context and for a purpose. Therefore, both digitization and the purposes of data access are part of the ethical landscape. After all, data sharing is a regulatory burden, so it should serve its purposes. Data access as regulation could become is a gatekeeper to the profession. Regulation always favors some participants in an enterprise and not others. Another way of conceiving ethics shifts from the individual level, focused on the obligations between researchers and research subjects or collaborators, to considering data sharing as regulation and purposes in the academy now.

This article argues that understanding data access as regulation requires first considering replication, the science-based justification for data access, in the context of widespread digitization. Next, this article reconsiders these debates about data access, accountability, and research transparency as regulation, including as audit. Audit checks work, which invites asking why and when work is checked. Sometimes checking works in conditions of mistrust. Checking may or may not address fundamental problems of mistrust. Therefore, the article concludes with an outline of mistrust in scholarship in the current work context. Data access as accountability to a principal could miss the mark if the point is to assure a broad public about the reliability of scholarship.

Revisiting the Replication Call: Verification, Reproducibility, and Inferential Replication

Data are often available via electronic archives and websites. Records to be transformed into data can be accessible whether or not scholars choose to make them available in data depositories. Lives online, including databases online, have blurred the distinction between work done at a distance and work done in person, with a close encounter between research subjects and researcher, or physical archives and a scholar. Data are digitized in at least three ways, useful to researchers of all epistemological approaches. First, people's online activities generate websites, videos, Facebook pages, pictures, and texts that all can be accessed and analyzed. Second, many records are digitized, including government records and historical archives (Lazer & Radford, 2017). Team projects, with both qualitative and quantitative data use, share data and analyses remotely. The online environment has changed the context for data access, influencing what data are useful, and for what purposes.

Advocates of replication argue that digitized data make replication available to everyone, as long as data sets include metadata and processes. However, original advocates of replicating in lab sciences as key to advances and trust focused on the processes, not the data. The science scholar Evelyn Ruppert argues (2015, pp. 128-129) that empirical replication was so difficult in the 17th century that good practice

for replication meant sharing the processes and trusting that they could work because someone had been appropriately trained. Writing down processes allowed sharing at a distance. However, sharing with interested people at a distance in the 17th century meant sharing with a much smaller group at a much slower speed than sharing in the online world. Scholarship today aspires to greater openness to more people than it did in the 17th century. Replication as depending on data access shared through digital archives reframes replication to sharing product, rather than the technology of sharing processes.

Gary King's (1995)[1] widely cited call for replication predated current problems revealed in embarrassing failures to replicate (Monaghan, 2013), scandals (Bhattacharjee, 2013) and well-publicized retractions (Retraction Watch, 2016). King argued that data access was essential to replication. Challenges built into the multiple meanings of replication suggest that replication as accountability for scholarship cannot be the only purpose of data access. When King called for making replication data sets available in political science, he focused on justifications from within the discipline rather than reliability for external purposes. King argued that replicability would allow scholarship to advance, which would be valuable to the field. Others would cite replicable scholarship more, he argued, and that gaining citations would motivate individual scholars. The principal was not the broader public, but *s*cience. Encouraging scholars by arguing that they would get more citations would line up individual incentives with the rather distant accountability to science. Furthermore, he argued that data access would allow scholars to build on existing work rather than beginning again. Deeper understandings of political phenomena require data access for replication.

Sharing records scholars produce across a distance with an unpredictable group of people (and perhaps no one, if no one is interested) requires trust without building it in familiar ways. Trust across a digital distance does not rest in getting to know each other or the family tree or pedigree of training. Sharing at a distance can open up research to more people and change gatekeeping, but it does not make trust easy. Career incentives are to catch mistakes or fraud by high-profile scholars. Taking down notable scholars may advance science, but the threat would not make it easy to share data. The very argument King made for data access and replication, that both would facilitate improvements in theorizing, disappears if freely shared; polished data sets are the products to be shared. The expertise required is data analytics on finished data sets rather than either the work required to make the world into data, or social science theorizing to ask good questions or to interpret (Mah, 2016; Wallach, 2014).

King wrote of the virtues of replication in political science before scholars focused on disaggregating replication into components. Replicability includes verification, reproduction (with another sample, using the same analytical processes), and inferential reproducibility (Tsai et al., 2016, p. 192). Verification implies that the same analysis will yield the same results, requiring methods reproducibility. Using the same methods is often harder than it sounds, since researchers cannot always know every method that mattered. Some processes leading to a result may be opaque, or a matter of fortune. King also argued that replication requires writing down all the research processes. As Camfield argues in this special issue, incentives in the profession work

against giving detailed accounts of research practices. In quantitative analyses, sharing all processes requires sharing the code to ensure that the same code used on the same data will yield the same result. Since code is also a creative work product, sharing computer code can be difficult for people who do not want to give away what took them a long time to create. Not only is empirical replication and sharing work processes or code difficult. A scholar of computational sciences who analyzes replication, Victoria Stodden, argues that empirical replicability is both difficult to accomplish and less interesting than the conceptually much more challenging inferential reproducibility. Few would argue that qualitative analytical methods that move between theorizing and data, and iteratively broaden the context for data, aspire to yield empirical verification (Charmaz, 2014).

Analyses still require imagination and interpretation, which means that replicability requires explaining inferential steps. These steps allow inferential reproducibility, meaning scholars draw the same conclusions from studies (Goodman et al. quoted in Tsai et al., 2016, p. 193). Data access need not lead to inferential reproducibility, since inferential reproducibility also requires laying bare inferential strategies, which are neither data nor work processes. As the psychiatry professor Alexander Tsai et al. (2016) argue, citing others, data sharing will not address replication concerns once verification (also called methods reproducibility), reproduction (taking a study to a different site), and inferential reproduction are disaggregated. As Stodden (2015b) argues, inferential reproducibility requires an ability to transcribe how someone thinks. More interesting than empirical reproducibility, Stodden argues, are computational and statistical reproducibility (Stodden, 2015a). Computational social scientists' creative product, the code, can be as difficult to share freely as any other creative product.

The experimental psychologist Adam Alter has argued against what he calls "replication as a safety net" even in fields where replication is the standard for data access. The possibility someone *could* replicate (even if no one ever does) leads people to publish fragile results from ill-designed experiments (Alter, 2015). Instead, training in good research practices alongside reinforcement for following what one has learned could improve work. In addition, replication by journal editors or others using a researcher's data set is not empirical replication, since the experiment is not redone, but computational and inferential. Neither can address empirical replicability and neither pretends to. Therefore, qualitative work is not distinctive in not having empirical replicability as an element of research accountability. Empirical replicability can be difficult for all kinds of sciences, and the threat is an anemic form of regulation.

Social psychology more than political science has undergone scrutiny concerning how difficult replication is, even for the quantitative work that researchers often believe is easier to replicate. Critiques both of experimental design and inferences from small samples have multiplied. Statisticians have critiqued conclusions drawn from experiments in social psychology with small numbers of participants. Looking for results that meet the standard of statistical significance, or p-hacking, is common research practice, promoting the fragile results Alter bemoans. The American Statistical Association issued a statement against reifying *p* values (Wasserstein & Lazar, 2016).

High-profile studies have proven difficult to replicate (Retraction Watch, 2016), and in a notorious case, outright fabrication brought down one social psychologist (Bhattacharjee, 2013). Scandal in political science has come through alarms over misconduct through investigation, not through routine monitoring. Routine monitoring might prevent scandals, but only if data are not fabricated. Monitoring still must be worth the cost.

Finally, some experimental work, the work that has most been subject to calls for replication, aspires to be useful outside lab settings. Replication in labs might answer the demands of science without meaning the results come about as intended outside experiments. Even the insights of research meant to be practical, such as behavioral economics, prove difficult to put into practice (Carroll, 2017). The difficulty of using the insights in the real world does not make them wrong. Conversely, the rightness of the scholarship, even its replicability in the sense of verification, does not make the scholarship serve intended principals beyond the science in controlled settings.

Even without sharing a spreadsheet, research at a distance is increasingly possible. Digitization raises new possibilities of working at a computer screen, including when conducting what once would have seemed to be archival or local work that required travel. With so much available on a screen, the use of information far exceeds the purpose of replication. Lives lived online mean that close interpretive work can be done with already digitized records. Digitization makes it possible to aesthetically appreciate records once only seen in physical archives. Records, lives, and projects already online alter arguments about data access, if arguments rest on an epistemological assumption about the importance of close work between researchers and subjects as fundamental to all qualitative research. Data archives include newspapers, documents aggregated for particular historical fields, property and family records on government websites, and reports from governments and NGOs. Some records available online are prepared as data, including about institutions unfamiliar to many nonacademics but of ever more evident significance, including, for example, courts, and the interest groups, business and individuals who bring suit (see, e.g., Cichowski & Chrun, 2017). Other records, including traces of a digital life, need to be turned into data, to be analyzed qualitatively or quantitatively.

Sharing information digitally, which once could only be accessed through travel, can allow people to see the world even where replication is irrelevant, sometimes thanks to work by scholars. For example, the legal historian Kimberly Welch has photographed court records and gravestones from Louisiana and Mississippi.[2] Welch has also photographed court records and the cemetery tomb of an interracial family. She explains the family lived openly with their mixed-race daughters. They pursued their interests in court in the Deep South before the Civil War. Although not traditional human subjects research, the work probably makes it possible to discern the history of people who are living. Although public records have always made it possible to link history to living people, records available on the Internet are available to anyone with an Internet connection, not only to those who would have traveled to look through decaying local government files. Welch has not only preserved records crucial to American history that no local government had any reason to preserve. She also made

Figure 1. Local court records in Claiborne County, Mississippi.
Note. From Kimberly Welch: http://www.kimberlywelch.net/images.html.

some of them widely available. Many African Americans fled the South in the 20th century's great migration. Families whose ancestors lived in 19th century Louisiana and Mississippi can access family history, even from far away. The county in Mississippi had no reason to preserve the records, especially by digitizing. Beyond the meaning to the family, the records upend some beliefs about slavery, separation of people by race, and African American's agency. Online access also allows a new researcher to draw on dimensions not of Welch's focus (see Figures 1 and 2; used with permission of Kimberly Welch). Similarly, anyone can use documents from Princeton's slavery archive (Princeton Slavery Project; https://slavery.princeton.edu). The records are invaluable to the history of family, race, and enslavement in the United States. The records could also be valuable for the very popular American pursuit of genealogy. It is impossible to know ahead of time who appreciates the accessibility of family records.

People who read records are likely to read them for different purposes and contextualize them in different ways. Making records available does not dictate a methodological approach, either quantitative or qualitative. Research at a distance already happens, and use can be for something other than systematic or generalizable scholarship or replication. The expertise of working at computer terminals already complements field expertise. The context for data access and research transparency changes the work and conditions for work. Work at a distance is different work than work done in the field, or when the field is a computer screen. Digitization is already here. Although it may be impossible to stop digitizing information, thinking through what digitization means about work, expertise, and theorizing still describes how work practices change. Displacement of expertise from fields to screens contributes to reorganizing the professional field. Research online can change what it means to work interpretively with close attention to context. With the expansion of lives lived on line, contexts can be virtual. Bits of information mean something different situated on one website rather than another. To know that information is increasingly available on a screen says little about who uses it, or how, or what the value is.
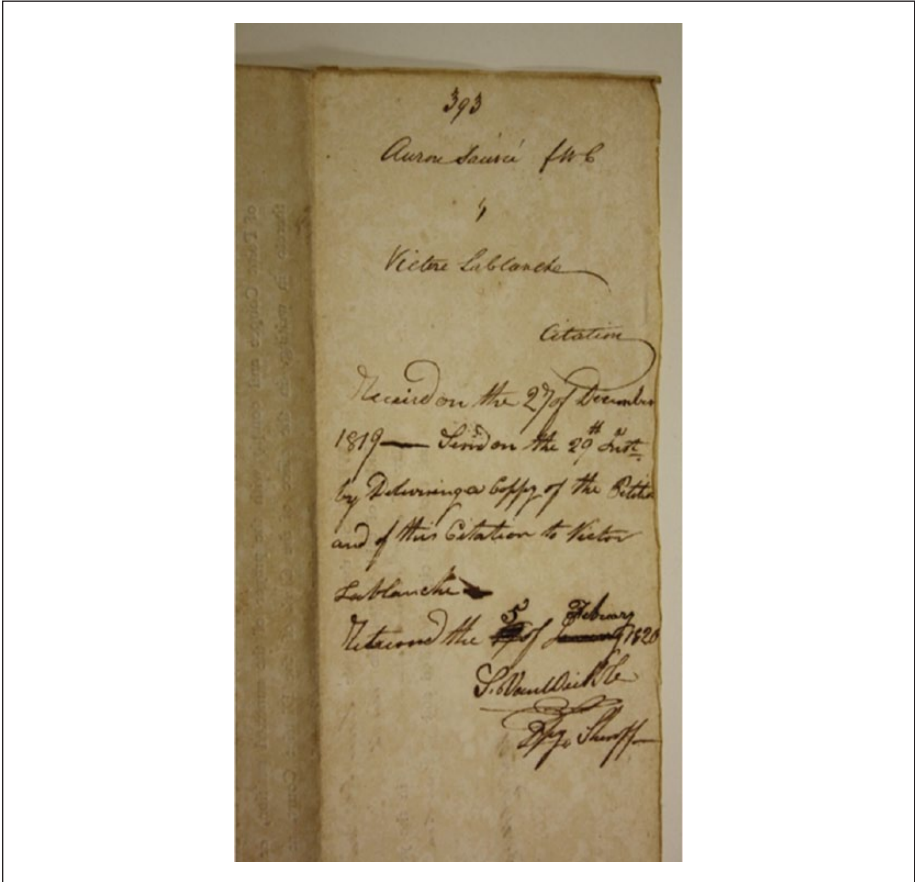
Figure 2. Photograph of a court record.
Note. From Kimberly Welch: http://www.kimberlywelch.net/images.html.

Replication once focused on shared research processes and calls for replication often still do. Digitization and new critiques of research procedures focus the call on sharing data sets. Replication based in a product shifts the accountability and learning from the process of research design, data collection, and analysis. Sharing products asks people to trust in those using their work product at a distance far removed from conditions building trust. Processes themselves, including the code that people write for quantitative analysis, are also creative products. Therefore, code and other products and processes for research easily used at a computer screen are potentially costly to share in the service of advancing knowledge. Alternatively, the work product need not be useful for checking work by replication. Instead, an intermediary product to be used for new purposes, such as Welch's photographs, can make more work possible, including work that exceeds the interests of scholarship traditionally understood.

Sharing images can imply the difficulty of work practices, but replication may not be the point. Welch's study of African Americans depiction of court records in Mississippi in the 19th century illustrates that sharing can contribute to seeing process. Her photographs give a rich feel of the materiality of working with court records in archives. The handwriting, the crumbling paper, the organization in boxes tells about the conditions of doing history in physical archives. Furthermore, Welch brought creativity and imagination to the study, drawing her own distinctive inferences. Writing down practices shared science at a distance in the 17th century. Now people can share with millions across vast distances, instantly.

The science-centered justification for sharing has been replication. However, other justifications weave their way through public statements. In addition to replication, calls have been to share to allow greater use of data in teaching, or use by a broader public, including students, journalists, and citizens. Since the profession does not control the work done with arrest records (Lageson, 2017), or other documents replication is only one way that data might get reused. Debates over the worth or possibility of replication miss all the other uses, some of which raise compelling ethical questions.

Making data accessible has promised to support replication. However, replication includes multiple elements, making it contestable and hard, even in favorable circumstances. Incentives often work against sharing all work practices, leading to sharing the more truncated work product. Submitting data and research practices to archives sets up distant, diffuse accountability to science. Sometimes payoffs are more immediate than advancing science, such as advancing careers. Incentives and institutional reinforcement of those incentives, including celebrity attendant on dramatic results, as well as changes in standards for research both encourage at least biased reporting of results and, sometimes, fraud. Insisting on replicable research in the face of acknowledged challenges with, at minimum, inferential reproducibility, will favor research with data that are easily shared and reinterpreted, thereby favoring some kinds of research over others.

Replicability as Accountability

Science is an abstract principal. Few know how to answer to the call of distant, impersonal science. Publication decisions by journals can serve as a more proximate enforcer. Many journals have data access boxes to check, and one in political science has mandated replication (discussed below). Failure to replicate can serve as shaming; the shaming puts costs of regulation on one individual when the goal is to improve all of science. That regulatory mechanism makes the failure personal, but research happens in a context of career pressure and payoffs for dramatic findings.[3] Shaming, compliance checkboxes, and rule enforcement against the backdrop of a larger goal are all forms of regulation. When the goal is difficult to achieve or hard to measure—advancing science—compliance with more proximate steps stands in for the goal, though the connection between the two may be uncertain (Haines, 2011). Compliance with rules deflects responsibility for accountability to the more contested long-term goal.

Accountability via data access for replication requires explaining how research products are selected for replication. They are not selected randomly, or for payoff for improving science. If selecting scholarship where reputational payoffs for demonstrating a failure to replicate are potentially high, then routinely making data accessible will not lead to replication. Although data output, code, and written down processes promise replicability, personal authority still holds sway in the academy. Reported misconduct in recent years attests to the continued pull of personal authority. The eminent experimental political scientist Donald Green took coauthorship on an article with fabricated data. Once he realized the junior author fabricated the data, he retracted the article. Replicating with fabricated data does not check work practices (Singal, 2015; Van Noorden, 2016).

An alternative to relying on the vagaries of checking spectacular claims or making a career of debunking a field is checking everything. Journals mandating replication before publication, with exceptions for proprietary or difficult to get data, make data access into a gatekeeper for careers. Again, perhaps that is the choice the profession means to make, but some cost-and-benefit accounting would help. In political science, the *American Journal of Political Science* (*AJPS*) has instituted mandatory replication by an independent group for each article that has been conditionally accepted (Jacoby, 2015).[4] Outside such unusual policies, replication may be more likely to work as a check when someone suspects a problem rather than as routine monitoring. Checking everything is a big commitment compared with spot checks, or going for high-value targets. In most regulation it is unusual to check everything. Contrasting styles of checking even have names in political science: police patrols or fire alarms (McCubbins & Schwartz, 1984). Responding to fire alarms searches out targets that are likely to pay off, also a strategy in tax audits. The promise to check everything is unusual and not usually recommended as a valuable use of scarce enforcement resources. The Institute charged with replicating for *AJPS* knows the demands of the work processes for replication. The cost—time for those checking—is also a benefit if it contributes to students' training. Aiming for a "total check of everything" aims for a stability of knowledge once sought out in the processes of peer review and competition over advances in knowledge. By not knowing whether anyone uses data, or for what, it is impossible to know whether data access serves the intended purpose of replication. However, checking every manuscript answers Alter's critique that the threat of replication is an ineffective regulatory mechanism. The future of research, or work that gets recognized as research, could become experiments that labs redo automatically to check for reproducibility (Wykstra, 2016). That will reshape work.

Replication, one motivator for calling for data access, proves difficult, even in studies designed as experiments in a lab setting, under variably controlled conditions. King's (1995) frequently cited call for replication justified it in terms of advancing science and advancing individual careers. More recently, replication has served as an accountability check for honesty in the profession. That turns to seeing data access as an audit of professional work. The language of audit as professional check first came as checks from outside; data access as a method implies internal checks as well as

sharing outside the academy for myriad uses. The article next turns to discussing audit.

Data Access as a Form of Audit

Digitization means that the era of research from screens, qualitative, or quantitative, is here. If data access is a gatekeeper to careers, mandating will change the work people do, scholars and others. Working with already digitized resources simplifies the challenges of contributing new data, including transforming work into digitized records. Checks are targeted, chasing celebrity and fraud. Clearn from omprehensive checks use substantial resources. Excellent research practices may sometimes make checking data unnecessary; it's hard to know. The multiple scandals in behavioral sciences involving prominent scholars and studies at least make it evident that scrutinizing data in celebrated studies will yield results. Neither comprehensive checks nor chasing celebrity tracks good regulatory processes in other circumstances, unless "celebrity" stands in for pursuing high-value targets in enforcement. Accepting that replication is difficult and full release of sensitive data is impossible, Alexander Tsai et al. (2016) argue that interview or field data and qualitative analysis allow for "audit trails." Exporting coding queries from qualitative data analytics software allows others to check interpretations. Scholars can share memos if not field notes. Analysis of the outputs of qualitative software affords advantages to scholars who have enough support to allow the use of software. Possibly even more significant in an era of falling prices for software, supporting reanalysis also requires support to manage the software, the memos, and the codes. Tsai and others' language of an audit trail for the outputs of qualitative analysis evokes audit culture, as well as the changes in expertise auditing supports.

Audit as a form of regulation is a useful way to think about data access. As the anthropologist Marilyn Strathern argues, audits impose a control with a perspective, while not recognizing that an audit comes from a particular point of view (Strathern, 2000). Audit as a form of regulation makes sense of the language of compliance and monitoring from previous sections of this article. As Strathern (2000) has argued, following the accounting scholar Michael Power (1997), auditing was initially for financial systems. Control is not imposed by daily supervision of work processes, but by technologies of monitoring (Strathern, 2000). In data sharing, the monitoring technology is that of reports, archives, and the expertise and funding for those who build both. Since audit checks the work product, the product of an audit then leads to questions about or affirmation of work process. If the data do not yield the results expected on reanalysis, or if the data look too good, suspicion then leads to questions about the process (Bhattacharjee, 2013). How well audit governs, at what cost, bringing what changes in expertise, are all empirical questions. Interpreting practices concerning government accountability in the United Kingdom, Strathern, Power, and others unpack auditing as a particular form of accountability. Although they focus on governing institutions other than the scientific enterprise, their findings are illuminating.

Figure 3. Map, reprinted in SFist (http://sfist.com/2009/01/09/mash-up_map_of_google_maps_and_prop.php).

Bracketing off the demand for accountability through audit as different, and more easily met, when large public databases are used turns away from the shared work contexts across quantitative and qualitative work. Those who focus on data that were sensitive as they were collected—interviews with members of vulnerable populations—argue that the accountability through audit is not worth the cost of exposing people who could not consent, which is not an issue in larger scale, homogeneous data. However, the process of analysis itself can make data that were not sensitive into data that reveal personal information (Metcalf & Crawford, 2016). Some records can reveal individually identifying information when users can merge or overlay databases on each other, especially at the municipal level.

For example, donors to California's anti–same sex marriage Proposition 8 found their homes revealed on maps. The public (by law) donor list mashed with Google Maps revealed where donors lived. People creatively relying on public records and computing capability resulted in widely available maps identifying individual donors and where they lived (Figure 3, taken from Keeling, 2009).

The privacy problems come in merging data sets not themselves protected; the problem of protection exceeds the sensitive interview data scholars sometimes imagine. Furthermore, since many people other than scholars can use databases, the problems enacted through data publicity and audit for the profession, and the profession only, misses the privacy problems raised by what data journalists or others might be

doing. Mistrust of the profession and its claims are answered with data sharing, but ethical problems have no clear governance for others who use records.

Not only do career incentives promote work that does not check out. Verification always has limits; "checking itself requires trust," as Power (1997, p. 2) argues. Illustrations above pointed out the ongoing pull of personal authority. Strathern argues that by definition, auditing is something imposed from outside. Even if someone can confirm an analysis through checking on field notes or running an analysis again—that is, the output of processes—the field notes or other data could be fabricated. Therefore, audit does not rely on the expertise of those conducting the initial work with whatever kind of data however analyzed, but on expertise in replication, or expertise in producing or analyzing audit trails from software. Expertise required to check can be different from expertise in original research. This expertise can displace research design, implicit knowledge, complex metrics, forming data through fieldwork, or imaginative use of administrative records and reports available online. Expertise through auditing outputs also accepts that rather than developing trust through learning processes of research together, trust develops through reanalysis.

Audit is especially suitable to research accountability where trust through learning to share good work practices is not possible, not least because sharing happens across wide differences. Using the product rather than the process as a way of understanding accountability accommodates the impossibility of writing down all the processes that led to the research outcome. As noted previously and as Camfield discusses in this special issue, the profession does not reward telling all the processes of research. She observes that incentives work against writing down all the research practices, since research seldom progresses as textbooks say it will.

Accountability through revisiting field notes, as Tsai et al. recommend, is not new. Checking field notes and triangulating reports allowed vigorous debate about the validity of the work of Margaret Mead (Shankman, 2009), decades after she completed her work. The imagined timeline for accountability through checking research outputs is shorter now in the era of digitization than in the decades passing before others worked through Mead's fieldwork or notes. Now, the *AJPS* replicates everything, as mentioned earlier, and journals at the time of submission can ask for affirmation that the data are available, or that data are exempt, or that authors describe the processes for getting access to protected data that cannot be released.

The possibility of checking memos and syntheses of field notes highlights that the call for data access for replication is only one of many possible accountability measures for the profession. Other gatekeepers in the profession are potential principals. These principals include employers who measure scholarly output. Measures of rapid output or work rapidly recognized as valuable could influence the kind of work done even more than demands for data access and create incentives to do something other than build the best science. Accountability to these principals creates incentives to do something other than produce the best scholarship. Some advocates of data access and research transparency acknowledge that one called data access "a threat that might keep potential cheaters honest" (Hamemesh quoted in Tsai et al., 2016, p. 192). A

check on how people represent themselves online in political science shows frequent misstatements of accomplishments (Van Noorden, 2015). Lessons from regulation are helpful: Breaches occur even in the presence of threats, and rich studies of regulation argue that more short-term payoffs, cultures, and contextual pressures all contribute to rule breaking or ethical lapses. Accountability to employers for production and the publication choices journals make, rather than science as knowledge, could explain some questionable decisions scholars have made. Furthermore, reflecting on how to evaluate whether the cost of monitoring is worth the benefit to improving scholarship would illuminate how effort in the academy is changing. Since King wrote in 1995, purposes of data access as regulation under suspicion rather than to advance either scientific knowledge or individual careers have multiplied, including checking on fraud, or revising how to conduct valid research (Bhattacharjee, 2013; Retraction Watch, 2016).

In publicly funded research, the public paying for scholarship is the principal. The reason the public would want the data, or which public would want data, are both hard to know, and worth finding out. If the principal is the funder and not only the distant demands of science, then it is worth asking what the principal wants. Public accountability from NSF includes the demands of science, since advancing science is its mission. However, public accountability includes Congress. If U.S. Congressional reports are a measure of the data that principals want, accountability for spending would have less to do with data availability and more to do with how well research can be explained and how inclined people are to believe in the value of the scholarship (Flake, 2015). Answers concerning accountability are multiple and do not rely on replication.

The principal for research accountability through data access may be undefined or multiple. Data access could be valuable for multiple reasons. If one purpose is audit, data access still leaves open the question of how to reach the purposes of auditing via data access. Once auditing is external checking of products rather than regulating work processes, professionals can still assess what work products should be available to those who might check. Auditing can be accomplished with the release of something other than field notes and interview transcripts (Tsai et al., 2016; Wutich & Barnard, 2016). However, audit cultures have costs. Costs may be part of the reason for objecting to data access.

Audit as accountability presumes an audit consumer. Yet audits are produced without a clear consumer; Power (1997, p. 2) writes that he never met a shareholder who wanted the product of his auditing work, though shareholders were the people audits were to protect. Thoughtful answers to the question of who would use a database would clarify the work that needs to be done. One check has found that seldom does anyone use publicly available scientific databases (Borgman as cited in Strasser & Edwards, 2017, p. 340). That is only one study; the production of data far exceeds knowing about use. Some use is for students learning analytical skills. Surely that is valuable, but it is not the same as ensuring scientific progress through replication. Changes in scientific theorizing have historically made data useless, as scientific paradigms change (Aranova, von Oertzen, & Sepkoski, 2017). For example,

data concerning phrenology and crime are useful only for historical interest, not for advancing understandings of crime since no one credible believes that head shape is related to the commission of crimes. Indeed, forgetting why such a study is worthless, or never knowing why, allows new, computer science–driven studies of faces for identifying criminals (Sullivan, 2016) or gay people (H. Murphy, 2017), efforts rapidly dismissed by social scientists. Graduate students might well reanalyze data in graduate classes around the country. Undergraduates learn through myriad data analytical strategies, sometimes using government databases. Better information on reuse would inform data access policies.

If a point of research transparency is not only reproducibility but also regulation via audit, data sharing is not an either/or proposition. Sharing coding queries from qualitative analysis software, or memos, or schedules of questions would answer concerns about the validity of scholarship without jeopardizing privacy or trying to achieve a potentially impossible replicability (Tsai et al., 2016). In a report from a U.S. NSF workshop on qualitative research methods, sociologist Susan Silbey argued that borrowing from the lab sciences, writing down research practices could answer calls for improving qualitative methods (Silbey in Ragin, Nagel, & White, 2004).

Making earlier products of analyses available incorporates the scholar framing questions and makes collecting data inextricable to the research process rather than displacing expertise. In responding to Tsai and his coauthors, Wutich and Barnard (2016) argue that documenting steps leading to inferences is well within traditional research practices. Documenting steps turns to inferential reproducibility rather than the empirical replicability that historians of science have pointed out is so difficult. Memos documenting practices can be shared, without threatening confidentiality, though, again, the profession does not reward sharing. King argued that sharing work practices is integral to data access. Sharing would contribute to accountability to the profession itself, if not empirical replicability.

Replication may be difficult, and audit may be a demand the profession makes of itself. Other people might still want data for some reason other than justifying the profession to itself. King's article focused internally: To build good science, the profession needs replication. Better scholarship for what purpose, though? The justification from NSF is that data the public has paid for should be available, for whatever purposes. That then also opens the question, who uses data? Is making it available allow people to do what they want, including pursuing hobbies of reanalysis, or appreciating maps, or genealogy? Or is it to build broader public trust in science?

Data access implies changing conditions of work by changing the value of taking time, for both quantitatively and qualitatively analyzed data. Data access comes at a time when universities rely increasingly on nontenure track faculty, who do not have the same responsibilities for research.

Conclusion: Changing Work Practices

In the changing academic climate, the value of slow work has dissipated. Principals for one's research are unclear and multiple, including one's university, the broader

public, the profession, and reliable and valid knowledge. Universities use metrics, and coaching encourages people to become more productive (Berg & Seeber, 2016; National Center for Faculty Development and Diversity; retrieved from https://www.facultydiversity.org/home). Slow work requires what Berg and Seeber in the *Slow Professor* describe as care and protection. The difficulty of replicating inferences, thought processes, and imaginative connections are not the only points of anxiety over data access and research transparency. Quantitative and qualitative scholars share concerns about the changing work environment, since rewards, recognition, and pressures for production are not exclusive to any one type of data or analysis. Valuing large numbers of rapidly produced publications with data treated as homogeneous and separate from its context, to be analyzed at a computer terminal by anyone with the expertise to operate computer programs with already written code, or memos about qualitative data analysis, fits with a changing academic climate. External funding at large institutions describes the work conditions of a shrinking proportion of scholars. Addressing the pragmatic concerns about data requires recognizing that problems of production of data, displacement of expertise, inference, context, and confidentiality plague large-scale data analytics too.

The audit culture that Strathern identifies for universities requires documenting work effort as well as competing for jobs and status on metrics in both the United States and the United Kingdom. Strathern developed her analysis of regulation by metrics of universities in the United Kingdom. Similarly, in the United States changing university contexts and reliance on metrics shape how data sharing and analysis in virtual space work. Metrics press scholars toward larger numbers of publications, and to game placement with higher impact factor journals. As discussed above, career incentives can exceed the power of accountability to the goal of advancing science. For example, difference in numbers of publications on the same issue contributes to the wage gap between men and women as a group (Leahey, 2007); under existing conditions, gaining more equitable salaries requires more repetition in high-profile publications, not more replicable scholarship. Pressure to rapidly produce multiple articles is only likely to increase for academics whose jobs include research. Increasing percentages of faculty are contingent, without time or a reward structure for research, let alone replication or data access.

Competition over what it means to be an expert, and the ability to displace expertise and upend academic authority, can be evident in both quantitative and qualitative work. The ability to work at a computer with information someone else put together implies that the professional authority accumulated through experience in a field, training, and the degree certification that the academy has valued diminishes. Replication with available data promises that the value of expertise in replicating will join the expertise of putting together data sets and conceptualizing a study if studies must replicate to be valuable. That promise can make everyone uneasy (Monaghan, 2013). This version of replication is very different from the first initiatives in the lab sciences, which relied on the processes of research.

Building and managing databases requires funding. It requires students who can transform notes or observations into data for a database. If data are not already

archived, it requires the skill and money required to load it into a maintained website. Implicitly, participating in the production of analyses that are on the cutting edge of data access contributes to the ongoing shift in what it takes to participate in scholarly conversations. Furthermore, how often people use existing databases is an empirical question. Surely it would be worth thinking about how often people use databases, for what purposes, and how mandates for digitization influence research.

Usefulness and trustworthiness to a world outside scholarship have become flash points in the United States since King wrote. Questioning science alongside the value of higher education pervades public debate, well beyond anything replication can answer. King did not address public mistrust of scholarship's value. His focus was on internal concerns for improving scholarship. Yet given the investment in the research enterprise and higher education in the United States and the widespread dissemination of bad information, scholarship's public value is at least as important as internal demands for replicability. When journalists coach people on how to make scholarship persuasive to a broader public, they do not discuss replication. Education in how to conduct scholarship is now often justified in terms of jobs, or the usefulness of scholarship in policymaking. The lack of reliability of scholarship in political science can rest as much in changes in the world as in processes never replicated. If one kind of research gets greater investment by private and public funders, resolution of regulatory principles will not matter as research practices change.

If the broader public is the principal to whom the scholarly agents are accountable, part of the point of data access may be to improve trust in science. Data access implies that trust in science must rest on data access that enables audit. It is an empirical question whether persuasion of the accuracy of results rests in audit or explaining processes of research. In analyses of climate change, persuasion in part rests on who is talking—one's friends or someone else (Kahan, 2012). Of course, communication studies themselves could be subject to audit. As Power (1997) has argued, trust is necessary to checking, since no institution can work without any trust. Across forms of analysis and data, the academy shares problems of public trust. Problems are myriad: among scholars who work at a distance from each other but can share outputs easily (Monaghan, 2013), among people who mistrust the sexual politics of scholarship, as demonstrated by sexual misconduct scandals since 2013 (Sterett, 2018), and between scientists and the public. Data access for audit may contribute to building trust among academics. It is an open question whether data access will contribute to building trust in the academy among a broader public that does not share either training or concerns about building scientific knowledge.

If data access is to foster public trust in science, it's worth comparing with other strategies. Urgent problems with regard to scholarship in the United States include explaining uncertainty, or gaining public trust in scholarship in the face of uncertainty. Data access and transparency is a wholly different kind of effort to build trust from ensuring that the research tells a compelling story, or that a trusted person communicates the findings. The argument for how data access makes scholarship accountable to its principal, the public, may be that scholars will check on other scholars' work, and scholars will trust outcomes better when they can check the data or critique work

processes. Trust will then filter to a broader public. Assumptions underlying that belief itself are worth exploring. Regulatory strategies relying on audit rather than supporting developing expertise in good practices, or learning how to conduct slow work, can stir unease under changing conditions in universities.

Notes

1. The same issue of the journal *PS: Political Science and Politics* with King's article included thoughtful reflections on the limits of replication. For a discussion of replication as regulation contemporaneous with Gary King's article on replication, see Fowler (1995). Those articles have not proven as influential in the debate over data access. The online environment and therefore the ease of sharing, if not the ease of crafting a database, have changed significantly since 1995.
2. http://www.kimberlywelch.net/images.html
3. Compliance can prove cursory in business regulation, and context matters (Haines, 2011). Looking at compliance from the point of view of the regulated makes sense of noncompliance (Gray & Silbey, 2011).
4. Mistrust has multiple sources. Mistrust of practices not discussed in debates over data access can spill over into mistrust concerning research integrity or the burden of replication policies. The spillover from complaints about unethical conduct is also central to the ethics of scholarship. Sexual misconduct as a source of mistrust in what counts as the best in scholarship came to *AJPS* in 2018, when the editor of the journal resigned after having posted a statement on the journal's website denying accusations of sexual misconduct (Gluckman, 2018). The Women's Caucus for Political Science linked this complaint about mistrust of the editor and of the association governing it to the low percentage of articles by women published in *AJPS*. The Women's Caucus letter also argued that the sexual misconduct could also account for possible reluctance on the part of women to submit to the journal. See Women's Caucus for Political Science, 2018.

References

Alter, A. (2015). Replication as a safety net. In J. Brockman (Ed.), *This idea must die: Scientific theories that are blocking progress* (pp. 373-375). New York, NY: Harper Perennial.

Ambrose, M. L. (2014). Lessons from the Avalanche of Numbers: Big data in historical perspective. *I/S Draft*. 25 August: 1-71.

Aranova, E., von Oertzen, C., & Sepkoski, D. (2017). Introduction: Historicizing big data. *Osiris*, *32*, 1-17.

Berg, M., & Seeber, B. K. (2016). *The slow professor: Challenging the culture of speed in the economy*. Toronto, Canada: University of Toronto Press.

Bhattacharjee, Y. (2013, April 26). The mind of a con man. *The New York Times Magazine*. Retrieved from http://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html?pagewanted=all

Carroll, A. (2017, November 6). Don't nudge me: The limits of behavioral economics in medicine. *The New York Times*. Retrieved from https://www.nytimes.com/2017/11/06/upshot/dont-nudge-me-the-limits-of-behavioral-economics-in-medicine.html?_r=0

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). London, England: Sage.

Cichowski, R., & Chrun, E. (2017). *European Court of Human Rights Database, Version 1.0 Release 2017*. Retrieved from http://depts.washington.edu/echrdb/

Flake, J. (2015). *Wastebook: The farce awakens*. Retrieved from https://www.flake.senate.gov/public/_cache/files/03714fa3-e01d-46a1-9c19-299553056741/wastebook—the-farce-awakens.pdf

Fowler, L. L. (1995). Replication as regulation. *PS: A journal of political science and politics*, *28*, 438-442.

Freese, J., & Peterson, D. (2017). Replication in social science. *Annual Review of Sociology*, *43*, 147-165.

Gluckman, N. (2018, April 18). An editor uses his journal's website to deny sexual-misconduct claims. Scholars revolt. *Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/An-Editor-Uses-His-Journal-s/243160?cid=wcontentlist_hp_latest

Gray, G., & Silbey, S. (2011). The other side of the compliance relationship. In C. Parker & V. Lehmann Nielson (Eds.), *Explaining compliance: Business responses to regulation* (pp. 123-138). Cheltenham, England: Edward Elgar.

Haines, F. (2011). Facing the compliance challenge: Hercules, Houdini or the Charge of the Light Brigade? In C. Parker & V. Lehmann Nielson (Eds.), *Explaining compliance: Business responses to regulation* (pp. 287-304). Cheltenham, England: Edward Elgar.

Jacoby, W. (2015). *American Journal of Political Science* Guidelines for Preparing Replication Files. March 25. Retrieved from https://ajpsblogging.files.wordpress.com/2015/03/ajps-guide-for-replic-materials-1-0.pdf

Kahan, D. (2012, August 15). Why we are poles apart on climate change. *Nature*. Retrieved from http://www.nature.com/news/why-we-are-poles-apart-on-climate-change-1.11166

Keeling, B. (2009). Map of Prop 8 Donors. *SFist*. January 9. Archived by WNYC. Retrieved from http://sfist.com/2009/01/09/mash-up_map_of_google_maps_and_prop.php

King, G. (1995). Replication, replication, replication. *PS: A Journal of Political Science and Politics*, *28*, 444-452.

Kulig, T., Pratt, T., & Cullen, F. T. (2016). Revisiting the Stanford Prison Experiment: A case study in organized skepticism. *Journal of Criminal Justice Education*. doi:10.1080/10511253.2016.1165855

Lageson, S. E. (2017). Crime data, the Internet, and free speech: An evolving legal consciousness. *Law & Society Review*, *51*, 8-41.

Lazer, D., & Radford, J. (2017). Data ex machina: Introduction to big data. *Annual Review of Sociology*, *43*, 19-39.

Leahey, E. (2007). Not by productivity alone: How visibility and specialization contribute to academic earnings. *American Sociological Review 72*(4): 533-61.

Mah, A. (2016). Environmental justice in the age of big data: Challenging toxic blind spots of voice, speed, and expertise. *Environmental Sociology*, *3*, 122-133.

McCubbins, M. D., & Schwartz, T. (1984). Congressional oversight overlooked: Police patrols versus fire alarms. *American Journal of Political Science, February*, 185-179.

Metcalf, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data and Society, January-June*, 1-14.

Monaghan, P. (2013, April 24). *They said at first they hadn't made a spreadsheet error when they had*. Retrieved from http://www.chronicle.com/article/UMass-Graduate-Student-Talks/138763

Murphy, A., & Jerolmack, C. (2016, March 19). *Ethnographic masking in an era of data transparency*. Retrieved from https://contexts.org/blog/ethnographic-masking-in-an-era-of-data-transparency-2/

Murphy, H. (2017, October 9). Why Stanford researchers tried to create a "gaydar" machine. *The New York Times*. Retrieved from https://www.nytimes.com/2017/10/09/science/stanford-sexual-orientation-study.html?_r=0

Power, M. (1997). *The Audit Society: Rituals of verification*. New York, NY: Oxford University Press.

Ragin, C., Nagel, J., & White, P. (2009). *Workshop on Scientific Foundations of Qualitative Research*. Washington, DC: National Science Foundation. Retrieved from http://www.nsf.gov/pubs/2004/nsf04219/nsf04219.pdf

Retraction Watch. (2016). *Yes, "Power Pose" Study was flawed but should not be retracted, says one author*. Retrieved from http://retractionwatch.com/2016/09/26/yes-power-pose-study-is-flawed-but-shouldnt-be-retracted-says-one-author/

Ruppert, E. (2015). Doing the transparent state: Open government data as performance indicators. In R. Rottenburg, S. E. Merry, S.-J. Park, & J. Mugler (Eds.), *A world of indicators: The making of governmental knowledge through quantification* (pp. 127-150). Cambridge, England: Cambridge University Press.

Shankman, P. (2009). *The trashing of Margaret Mead: Anatomy of an anthropological controversy*. Madison: University of Wisconsin Press.

Singal, J. (2015). *An interview with Donald Green, the co-author of the Faked Gay-Marriage Study*. Retrieved from http://nymag.com/scienceofus/2015/05/co-author-of-the-faked-study-speaks-out.html

Sterett, S. (2018). Law's presence, law's absence: Reporting employment discrimination in the academy. *Politics and Gender*. doi:10.1017/S1743923X18000119

Stodden, V. (2015a). Reproducibility. In J. Brockman (Ed.), *This idea must die: Scientific theories that are blocking progress* (pp. 529-532). New York, NY: Harper Perennial.

Stodden, V. (2015b, October 22). *Talk: Three kinds of scientific reproducibility*. Blacksburg: Virginia Tech.

Strasser, B. J., & Edwards, D. N. (2017). Big data is the answer . . . but what is the question? *Osiris*, *32*, 328-345.

Strathern, M. (Ed.). (2000). *Audit cultures: Anthropological studies in accountability, ethics, and the academy*. London, England: Routledge.

Sullivan, B. (2016, November 18). *A new program judges if you are a criminal from your facial features*. Retrieved from https://motherboard.vice.com/en_us/article/d7ykmw/new-program-decides-criminality-from-facial-features

Tsai, A. C., Kohrt, B., Matthews, L. T., Betancourt, T. S., Lee, J. K., Papachristos, A. V., . . . Dworkin, S. L. (2016). Promises and pitfalls of data sharing in qualitative research. *Social Science and Medicine*, *169*, 191-198.

U.K. Economic and Social Research Council. (n.d.). *Data management planning for ESRC researchers*. Retrieved from https://www.ukdataservice.ac.uk/manage-data/plan/dmp-esrc

Van Noorden, R. (2015, June 29). *Political science's problem with research ethics: Ethicist Trisha Phillips discusses the fall-out from the Michael LaCour scandal*. Retrieved from http://www.nature.com/news/political-science-s-problem-with-research-ethics-1.17866

Wallach, H. (2014, December 19). *Big data, machine learning and the social sciences: Fairness, accountability, transparency*. Retrieved from https://medium.com/@hannawallach/big-data-machine-learning-and-the-social-sciences-927a8e20460d#.iz0l4nf97

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: Context, process and purpose. *The American Statistician*, *70*, 129-133.

Women's Caucus for Political Science. (2018, March 29). *Letter*. Retrieved from https://womenscaucusforpoliticalsciencedotorg.files.wordpress.com/2018/04/wcps-letter-to-mpsa.pdf

Wutich, A., & Bernard, H. R. (2016). Sharing qualitative data and analysis. With whom and how widely? A response to "promises and pitfalls of data sharing in qualitative research." *Social Science & Medicine*, *169*, 199-200.

Wykstra, S. (2016, June 30). Can robots help solve the reproducibility crisis? *Slate*. Retrieved from http://www.slate.com/articles/technology/future_tense/2016/06/automating_lab_research_could_help_resolve_the_reproducibility_crisis.html

Author Biography

Susan M. Sterett is professor in the School of Public Policy at the University of Maryland, Baltimore County. She has published on accessing social welfare benefits after disaster in journals including in Law and Policy. Most recently she has published "Law's Presence, Law's Absence: reporting on employment and exclusion in the academy" in Politics and Gender.