ORIGINAL RESEARCH REPORT

Informal Laboratory Practices in Psychology

Jonna Brenninkmeijer, Maarten Derksen and Eric Rietzschel

Method sections in psychology articles differ in the amount of information they provide, or the level of specificity at which they do so. This can make incremental research (e.g., replication efforts) difficult, because potentially relevant methodological decisions or practices may not be reported. As yet, these unwritten practices have not been systematically studied; the current work represents a first attempt in this direction. For this paper, we interviewed 22 experimental psychologists in the Netherlands to find out about these unwritten aspects of experimenting, as well as their opinions and beliefs regarding these practices. Thematic analysis of the transcripts suggests that (a) experimental psychologists indeed employ a variety of methods that they consider relevant yet do not routinely report in method sections, and (b) these unwritten practices seem to fall into two broad categories: 'professionalism' and 'the production of good data'. We discuss implications for psychological research generally, and the replication debate more specifically.

Few discussions in psychology have been as widely publicized, as heated, and as lengthy as the debate on replication studies and the replicability of psychological research. One particularly interesting bone of contention in this discussion is the role of possible 'unknown moderators' in psychological research. Specifically, researchers seem to disagree with each other on (a) whether one may expect the method section of an experimental report to fully explicate the procedure required to replicate an experiment, and (b) whether failed replications (that is, studies that do not replicate an original result) may be (partly) due to these kinds of unobserved methodological or contextual differences.[1] These issues are important, because the replication debate in essence revolves around the reliability of psychological knowledge. As such, it is crucial to know whether variance in experimental results is due to error (i.e., random differences between data sets) or systematic variance (i.e., moderator variables). Moreover, because researchers often refer to these kinds of moderators when trying to explain why an experiment sometimes 'works' and sometimes does not, it seems likely that there is a class of research skills that future generations of psychologists could benefit from, but that are not routinely shared or transmitted through formal channels, such as method sections or methodology handbooks. The current article focuses on such informal

practices, as well as on researchers' beliefs regarding these practices.

Lab lore and informal practices

According to the American Psychological Association, method sections are generally required to describe "in detail how the study was conducted, including conceptual and operational definitions of the variables used in the study. (…) [A] complete description of the methods used enables the reader to evaluate the appropriateness of your methods and the validity of your results. It also permits experienced investigators to replicate the study" (APA, 2010, p. 29). In practice, however, method sections often seem to fall short of this requirement. This is partly due to space considerations (at least in paper journals), but several researchers have pointed out that it may actually be impossible to provide an exhaustive description of a study's method. Daniel Kahneman, for example, wrote that "this seemingly reasonable demand is rarely satisfied in psychology", because human behavior is too sensitive to contextual factors. "Experimental instructions are commonly paraphrased in the methods section, although their wording and even the font in which they are printed are known to be significant" (Kahneman, n.d.). Although not everybody agrees with this (e.g., Wilson, 2014), the notion that a 100% exhaustive method section is impossible seems to resonate with the experience of many researchers in Psychology. For example, Sanjay Srivastava, in one of his contributions to that same discussion, took a position in between those of Kahneman and Wilson: "In my thinking you can divide 'what matters' into 3 categories: the original researchers' specification of the experiment,

University of Groningen, NL

Corresponding author: Jonna Brenninkmeijer
(j.m.brenninkmeijer@rug.nl)

technical skills in the methods used, and common sense" (Srivastava, 2014). More is required for successfully doing an experiment than is specified in the method section. Srivastava did not explain what he meant by 'common sense', but he did note that the technical skills needed for the more complicated experiments (for example those that involve "confederates or cover stories") sometimes require special skills that are best learnt by visiting the original lab.

Thus, there may be skills and practices involved in the running of an experiment which possibly matter for the outcome, but that are not –or even cannot be– fully explicated. There may be more or less informal knowledge about these skills and practices. Srivastava (2016) has used the term "lab lore" in this context, for "knowledge about the pragmatics of running studies, passed down but not formalized". The term "laboratory lore" was also used in a series of papers in *The Behavior Analyst*, where it was defined as "that informal and miscellaneous collection of facts and assumptions concerning experimentation that is usually communicated in oral rather than written form" (Buskist & Johnston, 1988, p. 41). Buskist & Johnston called for research into the behavioral effects of these informal laboratory practices: knowing what influence, if any, they have on the outcome of experiments will allow the further sophistication of experimental methodology.

Research on lab lore and experimenter effects

There are several empirical studies that are relevant to this issue. Neil Friedman (1967), a student of Robert Rosenthal, noted that the experiment is a social situation, but it is not regulated as such in detail in either textbooks or methods sections. Such details as the distance between experimenter and subject, whether the experimenter may look or smile at the subject, the tone of voice or the speed with which the experimenter should read the instructions are not specified. Yet these aspects of the experimental situation, the "table manners of experimentation" as Friedman (1967, p. 73) called them, may very well influence the results. Evidence for this was provided by Friedman, who had different experimenters conduct the same simple experiment, formulated with the level of procedural detail that is typical in psychology, and videotaped the researchers' and participants' behavior. The experimenters' behavior varied considerably, and, as Friedman put it, "the differences made a difference; that is, what the experimenter did was a partial determinant of what the subject did" (Friedman, 1967, p. 52). Friedman predicted that replicability should increase "as experimenters become more aware of the role played by relevant contextual, scenic and interpersonal variables in the experiment itself and so write them up into the theoretical systems or frames of reference within which they work and think" (Friedman, 1967, 150). More recently, Stephen Gibson (2019) showed that much more went on in Milgram's obedience experiments than was described in the articles published about them. As the audio recordings of the experiments make clear, they had a rhetorical aspect, with some of the participants

engaging in argumentation and resisting the prods of the experimenter, and the latter often going beyond the script in order to persuade the participant to carry on.

Ethnographies in this area have focused on natural science labs (e.g. Collins, 1985), but more recently similar studies have been done in neuroscience and psychology, also identifying informal elements in laboratory practice. In his ethnographic study of brain imaging, for example, Simon Cohn noted that the interaction between experimenters and participants is vital to the success of the experiment, but is not explicated in the reports: "What is not acknowledged is the extent to which the volunteers were guided and directed, the physical arrangement of the scanning room, and the carefully orchestrated interactions between the participants" (Cohn, 2008, p. 95). Similarly, David Peterson described the importance of tacit knowledge in infant research. Infants are challenging experimental subjects, and proficiency in handling them is an important factor in the success of an experiment. "Some experimentalists display skill in creating stimuli that children find interesting. Others are good at keeping them happy and calm" (Peterson, 2016, p. 2).

Hidden moderators and adversarial collaborations

A classic example of the way in which 'hidden moderators' can affect research results was published by Latham, Erez, & Locke (1988), who conducted a series of studies to resolve their (previous) differences regarding the role of participation in goal setting. After deciding to engage in a collaboration to resolve the dispute, Latham and Erez discovered several possible causes for their differential findings, such as differences in the setting (e.g., dyadic versus group interactions), the framing of the experimental tasks (stressing their importance or not), and the wording of the instructions. Latham et al. point out: "Everything that an experimenter does in an experiment does not always appear in the published article. In discussions between Erez and Latham concerning possible reasons for the differences in their results, they discovered that the instructions the two of them typically used in the assigned goal condition were quite different" (p. 755). After a systematic series of studies, differences in instructions turned out to be the most important moderator of the effects. In the subsequent discussion section, the third author (Locke, who had acted as an impartial mediator in this collaboration) mentioned being struck by "the number of differences in procedure and design that can occur when two people are allegedly studying the same phenomenon", and pointed out that "If such differences occurred in these studies, one can assume that they also must occur in studies of other phenomena" (p. 769).

A more recent example of a similar kind of adversarial collaboration was published by Kerr, Ao, Hogg, and Zhang (2018), who studied the role of hidden moderators of the "minimal intergroup discrimination effect", again finding that (among others) differences in setting and instructions appeared to be crucial moderators of the effect. As the authors point out, these results were made possible by their doing "something that psychologists working on similar problems but in different labs and with different theoretical orientations rarely do—compare and

contrast our research protocols in close and very explicit detail, with an emphasis on identifying taken-for-granted procedures and methodological differences that might be psychologically important" (p. 67).

Although these examples suggest a role for hidden moderators and the importance of going beyond the information contained in the typical method section, the recent Many Labs 2 collaboration paints a different picture. This endeavour specifically aimed to replicate several psychological findings across multiple settings and samples. Results showed significant heterogeneity in effect sizes obtained, but this heterogeneity could not be explained by systematic differences between samples, settings, or procedures (e.g. the order of parts of the studies). The authors conclude that "dismissing failures to replicate as a consequence of such [hidden] moderators without conducting independent tests of the hypothesized moderators is unwise" (Klein et al., 2018, p. 77). Thus, the role of hidden moderators such as implicit, informal, or unwritten methodological practices remains contentious.

Although the possible influence of informal practices is a topic in discussions about unknown moderators in replication studies, what these practices themselves might be is seldom articulated. In this article we want to contribute to the literature about informal laboratory practices in psychological research by addressing this question. Specifically, our goal is to uncover such practices as used by experimental psychologists, as well as the beliefs and opinions these researchers hold about such practices (e.g., regarding their importance and effects). We present an explorative, qualitative empirical study, based on a series of interviews in which we asked psychologists about these unwritten aspects of experimenting. Our primary research question was simply: which informal practices, if any, do researchers recognize in their own work? Secondarily, we wanted to know what, if anything, researchers think about informal practices (e.g., regarding their effects and importance), particularly in light of the recent debates about the replicability of psychological research. After a section on methods, we present a thematic analysis of our interview material, organizing it around the main themes we identified in the transcripts.

Method

This article is the result of a collaborative research project between Theory and History of Psychology (MD and JB) and Organizational Psychology (ER) of the University of Groningen. MD and JB have a background in history of psychology and (ethnographic) science and technology studies and were familiar with the literature on informal research practices, while ER has a background in (quantitative) experimental psychology and had noticed informal practices first-hand.

Sample

After receiving approval from the Ethics Committee, we conducted 22 interviews with psychologists, from seven Dutch universities, in the period of March to June 2017.[2] We decided to limit our sample to social/organizational

and cognitive psychology in order to avoid practices that were specifically related to patients or children, and because both fields routinely use experimental research methods. The decision to conduct the interviews in the Netherlands was partly pragmatic, and partly related to the relatively strong integrity debate in this country – which implies that all participants were (more or less equally) familiar with the replicability debate and movements towards open science. 14 of our respondents were working in social or organizational (experimental) psychology, and 8 in cognitive or experimental psychology (see appendix).[3] To recruit these respondents, emails were sent to 54 psychologists who conducted (or had conducted) psychological 'lab' experiments, as we assessed mainly based upon their work as presented at the university's websites. 41 of these psychologists answered our email. Most of them were positive about our request for an interview, but not everyone had time or could be fit into our time schedule. Some psychologists declined because they did not see themselves as experimental psychologists (7), felt uncomfortable with the topic (1), or expected a fee for the interview (1).

Interviews and transcription

Most of the interviews were conducted in the respondent's university office, and two were taken by phone and Skype. All interviews were recorded, and transcribed *verbatim* by two master students from the University of Groningen, who were paid for their work but were otherwise not involved in the study. 20 interviews were conducted in Dutch and 2 in English. None of the transcripts were translated; the quotes used for this paper were translated and checked by the authors. The raw transcripts were sent to the interviewees for a last check and final informed consent, and except for some small changes regarding misspelled names or technical terms no changes were made. Because of privacy-related reasons, the transcripts were and are solely accessible by the principal researchers. For the same reason the interviewees have been given pseudonyms in this paper.

The interviews were semi-structured, meaning that all respondents talked about the same topics/questions, but not necessarily in the same order. We asked our interviewees about their research in the lab, and elaborated on this with questions about what they saw as important regarding the attitude towards participants, the instructions for participants, the design of the experiment, and the debriefing after an experiment. We also asked which research practices cannot be found in handbooks, and how their own research practices differed from those of colleagues, or from their own practices at the beginning of their research career. Furthermore, we asked which kinds of informal practices they would like to see described in handbooks. During the interview period, we had two moments of reflection in which we slightly adapted our interview strategy. The first adaptation was after the first 3 interviews conducted by MD and ER. In these interviews, MD and ER followed a specific order of questions while JB preferred to start with a general

question about the research of the interviewees and introduced her other topics/questions as the story of the interviewee proceeded, without a specific order. Halfway the interviews of JB, we added two interview questions to see if this would give any new information; one about the possibility to record experiments and one about the experience of failed searches for information for the set-up of an experiment. Both questions, however, were not very informative since our respondents had few relevant experiences with these situations.

To analyze the interviews, we used Atlas.ti7 and conducted a thematic analysis (Braun & Clarke 2006). Before we started coding, JB made a list of code words and explanations, and discussed these with MD and ER. Next she coded 3 interviews and these codes/quotes were again discussed with all authors. After agreement on code words, JB coded all interviews (except for one that was coded by MD). During this process, some codes were added, changed, and grouped. JB divided these codes (then 55)[4] into 4 broader themes. *Good science* (what is seen as good experimental psychology); *implicit practices* (which implicit practices and theories were mentioned); *transmission* (which practices were transmitted to students or colleagues and how); and *open science* (topics that were related to the open science movement, such as replication, sharing methods and data etc.). After this coding and grouping, MD did a second round of coding in which he checked all initial codes, and added or made some other changes in the coding; all in consultation with JB.

To analyze the material, JB made a first (semantic) inventory of all implicit practices that could be found in the transcripts. For this, she mainly relied on the quotes coded in the themes Good Science and Implicit Practices. MD used this first analysis to deduce some broader (latent) themes from these practices. Next, MD and JB discussed these themes, and returned to the material to find out in what other forms these themes could be abstracted from the material. In this back and forth between the authors and the material, the analysis in the next section was developed.

Results

After several rounds of coding we identified two general themes. We interpreted the informal practices that our respondents talked about as being driven by two main concerns, which we term 'professionalism' and 'the production of good data'. Professionalism, broadly speaking, refers to aspects of the way an experiment is presented to participants, as well as to the interaction between experimenter and participant. 'The production of good data' refers to those practices that are believed –by our respondents– to contribute to high-quality data. Beside these themes, we will also go into our interviewees' reflection on their informal practices.

Theme: Professionalism

The theme of professionalism connects several (informal) aspects of laboratory practice and seems to represent an important value guiding the work of researchers. For example, Albert said: "I'm not sure what's in the handbooks, but I do know what I find important.

[laughs] The most important rule of course is that you must approach [participants] professionally." Many interviewees mentioned this, and several (like Albert) talked about it at length. Several interviewees independently mentioned a professional attitude as typical of their own approach (as compared to others), saying things like "I think I care more about that than others" (Evert).

Taking care of the details

There are two, closely related, aspects to this professionalism. On a practical level, it means taking care of the details of the setting and the materials, avoiding errors or making a sloppy impression. The lab should be clean, tidy, and well-organized. It should have facilities such as separate and soundproof spaces, and a room temperature that can be regulated. It should not be chaotic, and there should not be any students chatting or hanging around. Distracting elements must be avoided. One of the interviewees gave the example of a lab with a poster of pop singer Madonna – "one of the biggest blasphemers you can think of" – which might have influenced the test results of Calvinists for example. In this way, researchers "add an extra stimulus to their design that no one will ever retrieve" (Simon). Another aspect that most researchers in our sample agree about is that the materials of an experiment have to look tidy, and that the formulations should be clear and short, without any spelling mistakes. Ursula, who uses paper forms, chooses to make a5 booklets instead of a4 forms, because (she thinks) it looks nicer and gives more of an idea of anonymity. George mentioned that he really does his best to make the lay-out of the material visually attractive, with pictures, fonts, and lay-out. Others, however, think this only distracts participants, and they try to keep the layout as simple as possible (Evert, Kevin), or aim at "just neatly" (e.g. Iris, Lars, Ursula), and "not ugly" (Kevin).

Respect for participants

Secondly, for our respondents professionalism is a matter of respect towards their participants. That respect is expressed partly in the practical orderliness of the lab and the experiment, but especially in the behavior, the bearing of the experimenter. The right attitude was often described as serious, business-like, but friendly. In other words, this second aspect of professionalism referred to adopting a certain 'experimenter role' and taking responsibility for a decent and serious interaction with the participant. The style of interaction with the participant is a very important issue for most interviewees. For example, they try to make their subjects feel comfortable (Clemens, Evert), and to show gratitude for their cooperation (Kevin). On the other hand, many researchers also warn against being too friendly (John, Paco, Simon, Ursula) or "amicable" (Hilda, Rebecca) towards the participants because this can influence the results. Nico, whose experiments involve taking saliva samples, wears a lab coat to convey authority and professionalism and thereby inspire trust, but also tries to be "nice" to participants, "so they don't think: who is this person bossing me around. Then you get resistance, you don't want that either." Many find a respectful and serious attitude important; businesslike

and professional. So, experimenters are not supposed to exchange personal details with their subjects (Evert). They should not engage in small talk (Nico), and there has to be a certain distance (Nico, Ton) between test person and the experimenter. Some researchers talked about a "reserved interaction" (Ton), others about "real interaction" (Kevin) or "just interaction" (Vera). Dora and Ursula use a standard welcome sentence (e.g. "Welcome to this experiment").

The importance of professionalism

Interestingly, and related to the second theme discussed below, several interviewees indicated that the professionalism of the experiment and the experimenter is important because it would make participants take their task more seriously, and thus yield better data. "If [the lab] is messy, yes, you see that immediately (in your data)" (Albert). Vera said that if you receive your participants with an air of indifference they will not be motivated to do their best, and you get bad data, although she had never actually tested that.

On the other hand, Frank and Lars, both cognitive psychologists, said that it does not really matter how researchers approach their participants: it is an ethical issue, but it has no influence on the results. To most interviewees, in fact, professionalism was above all an ethical imperative, not a methodological issue. "To me it is more a matter of decency towards your participant than that I think it's going to make much of a difference." Whether or not it leads to better data, researchers feel they must extend "a kind of professional courtesy" (Lars). Several interviewees connected this with informed consent and debriefing, but simply taking care of the paperwork was not enough: one should do justice to the reciprocity of the situation. "I think, that step that the participant has taken, you should reciprocate that." (Clemens) This involves more than just paying our subjects for their services. As Kevin put it, it is not a matter of "I pay you, you do this for me, and I don't have to see you again". Being professional, then, is part of the researchers' way of upholding their end of the bargain, of doing what researchers owe to the participant. The experiment is not just a social situation; as Friedman (1967) showed, the situation has a strong moral loading as well.

This connection between the material and the behavioral aspects of professionalism, and its fundamentally ethical nature, is expressed in a very Dutch word that was often used in this context, "netjes", which can mean neat, courteous, as well as decent. To our participants, a clean, well-running lab with a courteous experimenter appeared to be an ethical imperative and the essence of scientific professionalism in psychology. Researchers see their experiments as taking place in a social situation with a strong moral loading.

Theme: the production of good data

The second theme running through the responses to our questions is that informal laboratory practices in psychology are geared to a large extent towards the production of good data by facilitating and managing the performance of the participant. There is a theatrical quality to experimentation, with the participant as a somewhat fickle protagonist in a play written by the researcher and directed by the experimenter. Getting good data out of one's experiment requires not only a good script but also stagecraft and direction. There is overlap here with the theme of professionalism: good staging of the experiment is a matter of making sure the lab is tidy and clean, and that there are no distractions. As we discussed above, at least some researchers believe this gives better data.

Scripting the experiment

When we asked our respondents about the design of their experiments, their answers often touched on *scripting*. Many researchers said that the design has to be so clear that the participant knows exactly what to do. A good experiment, Albert told us, is an experiment that runs "as if it has been running for years". The order of events (e.g., measurements) is particularly important, but respondents had rather different ideas about which order should be used. Some interviewees said that it works best to start with the most important variables, because participants are more focused at the beginning of an experiment. Others, however, think that is better to start with demographic variables (gender, age, etc.) and then move on to the most important measurements. One researcher mentioned that she mixes positive and negative items to make the test a bit more unpredictable. Three interviewees mentioned that one should not put too much time (not more than 5 minutes) between the dependent and the independent variable. How much time an experiment is allowed to take, however, differs a lot according to our interviewees. Clemens said "The tasks are not nice. (…) I would prefer to continue for three hours because that gives me good data, but after an hour and a half my participants can't take any more." Others try to make their tasks nice and fun, and make sure they take no longer than 20 minutes (Hilda, Lars).

Directing and motivating participants

Apart from scripting and staging (creating a professional, distraction-free laboratory environment), the production of good data requires careful direction of the main performer, the participant. Directing the actors on the experiment's stage poses a special challenge: the participants are supposed to act naturally, be themselves, spontaneous and naive. Bringing out this natural performance in their participants is an important concern of researchers, requiring practices that are often not articulated in the method section. This starts with the instructions for the experiment. Several interviewees spontaneously noted that method sections seldom contain the full text of the instructions (Boris, Evert, Simon), yet everyone said that the instructions are a very important part of the experimental procedure. About the formulation of the instructions researchers often said that it has to be really foolproof. Instructions should be "long enough but no longer" (Wilma), "the fewer words you need the better" (Boris), but they should also be "absolutely crystal clear" (Kevin) and "dummy-proof" (Boris). In other words, a good instruction should be easy to understand without a lot of effort. The formulation is seen as very important because the success of the experiment depends

on a good instruction, not only because the participant needs to know what to do, but also because good, clear instructions motivate participants to do their best (Boris, Evert, George, Iris, Kevin). Motivation is important for all interviewees, but especially for the cognitive psychologists in our sample, whose experiments are often (according to them) rather repetitive and boring. "The majority of the experiments are tasks which I know from experience are boring to do" (Boris). Making sure that participants stay focused and take the task seriously is thought to be of great importance in these experiments, and our interviewees mentioned various tactics to motivate the subject. For example, Evert shows the average result or reaction time of the task and asks individual participants to make notes about their own reaction times or results. Frank and Lars emphasize at the start that the participant is very important for their research.

Mentally preparing the participant

Beside motivating the participants and ensuring that they know what to do, the instructions are also important to bring the participant in the right frame of mind; not just in terms of experimental manipulations, but also to make sure that participants are receptive to the experimental procedure itself. For example, if the experiment may involve discomfort for the participant, care must be taken to mentally prepare the participant for what is to come. Clemens, who works with Transcranial Magnetic Stimulation (TMS), prefers to err on the side of caution and exaggerates the discomfort of the procedure, so that participants will be relieved when it is better than expected. Three interviewees who work with fMRI talked at length about the instructions to participants. Rebecca said that making participants feel at ease requires "people skills", a lot of talking, and much time. Paco uses an extensive informed consent form, but because it takes a long time to read it and he wants to make sure that the participant understands he goes through the form together with the participant, elaborating on things where necessary.

Some researchers write more extensive instructions than others. One social psychologist says:

> I notice that people differ a lot, for example, when a questionnaire is introduced, some people just say: cross the number that applies, and other people give like a whole outline and give some background and try to get people into the situation and provide a context and, maybe even say: well, we're interested in such and such, so help them understand it a little bit, the background and the purpose of the questions. So, I think that's a practice that you don't necessarily see in textbooks, so how much introduction you give to certain questionnaires or tasks. (Wilma)

This is also what we noticed in the interviews. Boris, for example, first describes in general words what the experiment is more or less about, and then repeats step by step what will happen. Others prefer to use standard sentences. Wilma says she also tries to "keep instructions short", but tends "to give more":

> So, I also like to give headings so that you announce the topic and then you write a little bit of text and, so (...) I would tend to be rather extensive in explaining and saying things differently a few times, so enquire whether they understood it correctly.

That is, some researchers repeat the instructions between the tasks, or they include questions to see if the participants understood what they should do. Others, however, warn about too many instructions. If you give too many instructions, Lars said, you prime the participants how to respond. Nonetheless, Lars thought that writing good instructions is not difficult, but rather self-evident.

Not only should participants know what to do, often they must also be brought in a certain state of mind. Social psychologists in particular often want participants to imagine themselves outside of the lab. George for example said: "What I find important, I try to make it lively for the participants so that they can imagine themselves in the situation". When we asked how he does that, he answered "well, with the, yes, the instructions". Sometimes such instructions take the form of a straightforward assignment to participants to imagine themselves in a particular situation. "Imagine you are sitting in front of the television, alone", "look at it as if …" (Hilda), "imagine you really are outside on the beach playing a ball game with three other people" (Paco).

Increasing psychological realism

Two researchers augment the illusion of reality with deceptive actions by the experimenter. They described "little tricks you can use to make the situation convincing" (George). In both cases the experimental illusion involves the participant purportedly interacting with other participants (actually a computer program). To strengthen this illusion, George prefers scheduling the experimental session when there are other people present in or near the lab. Sometimes, he will pretend to be talking with people in another room, and will open and close doors to give the impression of other participants being shown into their cubicles. He called this "one of those implicit things that I have sometimes made explicit to students, but strictly speaking I don't know whether it actually makes a difference" (George). The other researcher, Paco, engages in similar pretense (opening doors, pretending to interact with other people), and he was not sure either whether it matters.

Tricks to mislead the participant were a somewhat touchy subject among our interviewees. Some interviewees from social and organizational psychology reject deception outright. Apart from moral considerations (Iris said "it's just not right, we don't lie to people") it is also a matter of maintaining the right frame of mind among (future) participants. If it becomes known that the lab uses deception, they reasoned, participants will no longer take the experiments seriously (Albert, Kevin, Nico, Oskar), and moreover honest information and real feedback gives a stronger, more realistic result (Iris, Kevin). Other researchers employ various tactics to secure the naivety of their test persons. Hilda never informs the participants completely about the experiment during the debriefing,

because they might later take part in similar experiments and they must not have "foreknowledge that will mess up the experiment". She acknowledged that this is "not quite according to the rules". Other researchers exclude participants who have already done a comparable experiment, or only use participants from the internet, outside of the university, or another department, because they think that psychology students are too experienced as participants. When researchers have to rely on psychology students –which often is the easiest way because Dutch students are obliged to participate in experiments– they generally have a preference for first-year students, because these are still the most naïve.

Formal and informal checks

To make sure that participants will behave, are behaving, or have behaved as they should researchers employ various kinds of checks. Of these tactics, manipulation checks are the most explicit, since it is standard practice to mention them in the method section. Ursula, however, noted that the method section will often contain only one example item of a manipulation check, leaving other researchers to guess about the wording of the other manipulation check items. Researchers also do more informal checks. It is common to ask, when the informed consent form has been signed, "is everything clear, do you have any questions?" (Paco). Albert reported that he does that especially in the first couple of experiments in a series, so that he can tweak the instructions if they turn out to be unclear.

To verify if the test person understood the task, researchers sometimes use reading or attention checks (was x a man or a woman?). Another tactic is to ask participants questions about the tasks (what do you have to do?) and if the subject fails, the instruction is repeated, and sometimes the question is repeated until the subject gives the right answer. Researchers can also use pauses in the presentation of the text that prevent subjects to continue, so that they will more or less be forced to read the instruction, and some interviewees ask test persons to report when they are ready with a task to check if they have spent enough time on a task. Frank uses warning signs that cannot be clicked away for a couple of seconds when the participant does something wrong or does not pay attention.

Although most respondents make use of such checks, George and Rebecca noted that checks, particularly manipulation checks, carry the risk of making the manipulation too salient, and therefore unnatural (or, depending on the order of the measurements, they might even put the manipulation at risk). Ton claimed that "we have sometimes destroyed effects because we asked people to summarize what the instruction was". One solution is to move the checks to the end of the experiment (Dora, George). Paco however stressed that he finds it important to "carefully explain the instructions, repeat the manipulations several times, so that they're really hit over the head with it". Simon said he often repeats the instructions after a break, and added that he had "never put that in a paper, so that is indeed kind of implicit".

To find out if the experiment worked well some researchers use checks at the end of the experiment. Hilda does an especially extensive debriefing after the first three experiments: "what did you think? Was everything clear? How did you experience it?" This researcher adapts the procedure for the rest of the series of experiments if the answers to these questions indicate problems. Others ask participants what they thought the experiment was about (Dora), whether they had noticed the deception (Paco, Nico), or more generally what they thought of the experiment (Boris). For Oskar, this last question was mainly a "service to participants", that is, a kind of service that participants received in exchange for their data.

Dual concerns

In summary, many of the informal practices that our interviewees talked about seem to serve the goal of producing good data by managing the performance of the participants: by providing a distraction-free environment, by making sure the events flow smoothly and the participants can focus on the task, by instructing the participants and putting them in the right frame of mind, and by checking them and correcting them if they go off-script. All this work is dominated by the need to be forceful, yet subtle and unobtrusive, because the performance of the participants has to be natural and not colored by the directions of the experimenter.[5] However, the theme of professionalism shows that our respondents' approach to their participants is not merely instrumental: they are acutely aware of the social and ethical aspects of the experimental situation, and actively and consciously take these into account in the way they set up the social and physical aspects of their research.

Interviewees' reflection on their informal practices

Since, as we noted at the start of this paper, the norm in psychology is that the procedure of an experiment is completely described in the method section of an experimental report, questions about informal practices are inevitably loaded. Informal practices are of uncertain legitimacy, and this has only become more pronounced in the current crisis of confidence in the discipline. A clear symptom of the sensitivity of the topic was the fact that a few interviewees seemed rather on their guard during the interview, and someone told us, after the interview had been conducted, that he had suspected us of looking for evidence of bad practices. We may have inadvertently caused this reaction ourselves, by assuring potential interviewees, when we first contacted them, that we were *not* after questionable research practices. There is much suspicion in the discipline at the moment.

During the interviews and while going through the transcripts, we felt that respondents' attitudes towards the topic and the interview in themselves represented an important result. Thus, beside summarizing the 'substantial' remarks about informal research practices, it seems worthwhile to also give an overview of interviewees' more reflective remarks about the issue at hand. That is, how do researchers see the issue of informal practices, and why do they think these practices are (and remain) informal to begin with?

Reactions to the topic

As mentioned above, raising the topic of informal practices, as we did in the interviews, often provoked reactions from interviewees which in one way or another relate to the normative issues regarding such practices, and the debates about them. A first example of this is that the way we introduced our topic to interviewees often led to some embarrassment. In the interviews, we explained that we were interested in skills and practices that researchers consider important in conducting experiments, but that are usually not explicated in handbooks or methods sections. Although all our interviewees eventually understood what kind of skills and practices we meant, their first reaction usually was that they do not read handbooks for this kind of information. They would laugh and say things like "gosh, well I'm not very familiar with what's exactly in those books" (Boris). When we asked whether such a handbook would be good to have, most reacted as if this were the first time they had considered the matter. Some were positive. One researcher, who had edited a handbook on the statistics of experimental research, thought that "a handbook about what's involved in doing good research would be very useful" (Frank; Hilda said something similar). Other interviewees jokingly suggested we should write that handbook ourselves (Wilma), or thought that a study like ours could make researchers more aware of how colleagues in other labs worked (Albert). A more common response, however, was that it would be very difficult (Dora, George) or even impossible (Ton) to compile such a handbook, because there are many things that may be important, and they differ per experiment. "There are so many little things that people do because they think they are important, that it's not doable really to put them together in a sort of catalogue" (George). Another practical problem that one respondent mentioned, was that internet and communication technology –increasingly important in conducting scientific research– are developing so quickly that such a handbook would quickly be hopelessly outdated (Lars).

Some interviewees, on the other hand, did not immediately accept the idea of informal research practices as such. When we asked Frank, who is prominent in the reform movement in psychology, whether he has "certain tricks" that are not in the method section he was adamant: "No. No, if they're so important, they're in the method section." He then added, however, "or implicitly, right". Oskar likewise first insisted that everything that matters is described in the method section of his papers, but then added "well, as much as I can". Lars, also a prominent reformer, was at first rather negative about the idea of informal practices, but towards the end of the interview did say "there are things you learn by doing and are never written down, I think". Such things should be "relatively small", however. Other interviewees were much more comfortable with the idea of informal practices. Both Evert and Nico said that "of course" there are things that are not written in handbooks and method sections. George had prepared for the interview by thinking about "what I've got for you" — the first thing that had come to mind was how to instruct and inform participants in such a way that everything is clear, but without saying "this is exactly my research question".

Why do some practices remain informal?

The question why practices are sometimes informal and unwritten was touched on by several interviewees. One reason that was proposed is that they may simply not be very relevant. Boris said that "side issues like how is the participant actually directed through the experiment – often that is not described." According to John, what is required in a paper is an "effective description (…) Some parts of the process are more important to mention for that than others. In the context of an article of a certain length." Researchers also thought that some things are simply "obvious" (Boris), not worth talking about or spelling out. George said: "you don't talk about it a lot, because it, this is not something you discuss a lot at conferences, because, well, it's not, it's not very exciting. I think a lot of people just do it like this."

On the other hand, and in line with the notion that some of these practices are in fact thought to contribute to high-quality research data, Albert, Boris, George, and Nico also said that there should be more discussion about informal practices. Researchers are only partially aware of how other labs work. They know, of course, how things go in the labs where they used to work, but when we asked them whether their practices differ from those of colleagues their answers were usually hedged with "I think…". As Nico put it: "You never discuss that because you don't tell each other, you don't know it about each other", adding that our project might change this. Maybe through a project like this, Albert said, researchers might learn from each other how to organize a lab or collect better data.

Among interviewees there was broad agreement that it is better that informal practices are made explicit, at least if they matter for the outcome of the experiment. The ideal of Open Science was often mentioned. In fact, the researchers who are most active in the Open Science movement or are committed to its ideals, like Boris, Frank, and Lars, believed that their methods sections are already complete, although, as mentioned, even these researchers eventually went on to give examples of unwritten aspects of doing (their) experiments. However, some interviewees expressed caution about making practices explicit, in particular when it takes the form of rules or guidelines. Dora, George, Oskar, and Ton warned that the best way to do an experiment depends very much on the particular study, the context, and the kind of participants. It is important, they said, to be precise and explicit in a particular study, but "there are so many differences between studies and designs and manipulations" that it is very difficult to write guidelines that cover all of them (Dora). Moreover, Dora added, the freedom of the researcher should not be limited too much.

Discussion

The picture that emerges from these interviews is one of simultaneous consistency and idiosyncrasy. Regardless of the decisions that researchers make, all our interviewees think (or have thought) intensively about most aspects of their research, and agree with each other on a number of things.

Consistency and differences

Professionalism, taking care of the experimental materials and environment, dealing fairly and decently with participants, making sure that participants remain motivated and attentive, preventing overload and confusion, checking whether participants understood and followed instructions, were all aspects that all or most respondents mentioned as important 'good research practices', even if the actual practices themselves (e.g., using a5 booklets, using headers and repetitions in the instructions, etc.) differed. Further, most participants agreed that method sections indeed do not contain all relevant information regarding details of a study, such as the actual instructions used.

In contrast, some interesting differences between respondents were their attitudes towards (mild) deception (ranging from 'not OK' to 'sometimes necessary'), the order of measurements (e.g., demographics first, manipulation checks immediately after manipulation or later on, etc.), the use of brief vs. elaborate instructions, and the degree to which experimental protocol should be standardized (as opposed to, for example, adapting them to the personal style of an experimenter). Many interviewees believe it is important to standardize communication with the participants as much as possible and therefore present information and instructions on paper or on screen, yet others think it is better to inform and instruct orally, precisely to make sure that every participant understands what is communicated in the same way. Another example: Dora, who likes to write very precise scripts for her experiments with exact formulations for the instructions, even adapts the wording of the instructions to particular experimenters, because "some words don't work for certain people"; none of the other interviewees was so precise in varying the script.

Many interviewees saw themselves as having a *personal style* of doing things. These were ways of doing that they referred to as "personal" (Iris), "my way" (Hilda, Simon), "everyone has their own way" (Frank), "my hobby horse" (Rebecca), or "a little personality trait" (Ursula). These personal ways varied widely, and could consist in visiting the lab to check on their PhD student's experiment (Albert), "overkill" in the number of participants (Frank), trying to make the experiment not too boring for the participant (Hilda), avoiding deception (Iris), or extreme orderliness (Ursula). Interestingly, many interviewees indicated that they considered these ways to be typical of their work, but did not know whether others worked this way too. Even professionalism, a value that was actually widely shared, was often considered idiosyncratic.

Evidence-based practice

Apart from these similarities and differences, however, a striking finding was that many researchers seemed to be aware that their beliefs and practices were not necessarily evidence-based; that is, many remarked that they did not actually *know* (e.g., through research) whether or not a particular practice actually 'worked'. Even if they had tried out different ways of doing research, the conclusion that a particular practice worked best seems to have been reached more or less informally. There is an interesting inconsistency here, in that researchers agree that it is important to do things 'right' if one wants high-quality data, but have rarely if ever actually put their implicit methodological beliefs to the (empirical) test (for example, by trying to falsify rather than confirm them). To a certain extent this is inevitable, since researchers simply cannot test every possible implicit belief if they want to get around to actually testing their substantive hypotheses. The classic methodological research on experimenter effects and biases does seem to have led to a broad awareness that all sorts of factors in an experiment can affect results, but this awareness partly appears to translate into idiosyncratic decisions on how to deal with those factors, rather than an ongoing line of empirical research that can systematically inform a new generation of researchers. In fact, a PsycInfo search on papers with the term 'experimenter effects' in the keywords (published between 1950 and 2018) yielded only 29 hits, 11 of which concerned parapsychological research – suggesting that research on experimenter effects is not exactly in the vanguard of present-day experimental psychology. Then again, the Many Labs project mentioned earlier does explicitly address the role of hidden moderators, even if not at the level of individual experimenters' behaviors. It may simply be that research on potential hidden moderators has taken on a different form that is more in line with the present-day focus on 'big data': rather than conducting experiments specifically designed to test the effects of certain (e.g., procedural) factors, researchers now collate datasets and test for possible moderators there.

Whether researchers' choices and styles actually make a difference cannot be decided on the basis of our material, of course. However, we can note two things. Firstly, many interviewees believe these are important choices, but do not always have evidence for their influence, and some researchers in fact acknowledged that they did not know whether this or that informal practice actually matters. This means that these practices are not only informal in the sense of 'not described in method sections', but also in the sense of not having been explicitly tested for effectiveness.

Secondly, it became clear that researchers in some cases do not *know* that other researchers do things differently, or, as in the case of professionalism, the same. Such patterns seem to be invisible to the researchers themselves, precisely because of the informality of the practices. Thus, beside deciding not to include such practices in formal method sections, even informal (e.g., personal) communication about these practices

between researchers who are not collaborating directly seems limited. One interesting question is whether competitive motives may play a role here: It is possible that researchers are more likely to informally share their individual 'tricks' with close colleagues (and collaborators) than with members of other, competing research groups. If this plays a role, one may wonder what the effect will be of the current move towards Open Science and the (mandatory or voluntary) sharing of research materials and data.[6] Creating a climate of openness and sharing can be difficult to reconcile with a system that stimulates and rewards competition, and this may mean that implicit practices are likely to stay implicit. Alternatively, some of our interviewees' statements also suggest that the possibility of discussing these issues simply does not occur to most researchers, and that they would actually be interested in doing so if they knew that other colleagues make different choices.

Finding the balance

Many aspects of informal laboratory practice that were mentioned involve striking a balance: the experimenter's bearing should be business-like but kind, one should inform participants about the goal of the study without giving the hypothesis away, the instructions should be clear but not directive, and manipulation checks must not make the manipulation too salient. Moreover, researchers' instrumental (high-quality data) and ethical (a pleasant and fair social exchange) goals need to be reconciled. All of this seems to require a measure of judgment that is the result of experience or plain common sense. When we asked our interviewees how they determined what was the right way to be professional or to inform and instruct participants, they usually did not give explicit and precise rules, but referred to intuition, feeling, and imagination. An ability to put oneself in the position of the participant is thought to be crucial (Boris, Kevin, Rebecca). For example, asked how she teaches students how to write instructions, Rebecca answered "uhh, [long pause], well, that they try to imagine how, how something like that comes across in a natural way; so, a bit like that".

How to find this balance is not very easy to explicate; it is tacit knowledge. As John explained: "Psychological experiments often ask for, for a balance between things. So, you have to sit relaxed, but also be concentrated. Your responses have to be accurate, but also fast. So, those are things, it's often easier just to show them." Thus, the informality of these practices, the fact that they are not precisely defined and written down, seems to be related to the fact that researchers need to strike the sweet spot between two extremes; apparently this 'Goldilocks' problem is solved in ways that defy precise definition. This would imply that at least some informal practices cannot be formalized, and thus will always defy the requirement of making the experimental procedure completely explicit. In other words, however elaborate and detailed a method section gets, there are always going to be some aspects of the reported study that are not included – which might, in theory, affect future (replication) research results. It is worth noting, however, that for some of our interviewees (Lars for example) such matters were a lot more clear and straightforward than for others.

Strengths, limitations, and future research

To our knowledge, this study is the first attempt to systematically address informal or implicit research practices within the domain of (experimental) psychology. Our use of interviews to do so naturally has advantages and drawbacks. The semi-structured format used in this study allows for some degree of standardization and comparison, while retaining enough flexibility to allow participants to address whatever aspect of the topic they find important. Of course, the information gleaned from these interviews depends strongly on the interviewees' motivation and ability to explicitly reflect on their informal practices. This is potentially problematic when studying practices that may, to some degree, be implicit (in the sense of not explicitly formulated) and hence may escape conscious reflection. We therefore set up our interviews in such a way that we invited interviewees to speak freely about the entire research procedure, from the recruitment to the moment the participants leave the lab, regularly probing for details that they might think too trivial to mention. We believe the results we have presented here show that interviews are a useful approach to the study of informal practices. Nevertheless, ethnographic observation would be a valuable complement to our study. Such a follow-up is necessary to find out if there are any differences between what people say and do, and may perhaps bring up other informal practices that are now overlooked.

This is one of the first studies of informal practices in psychological research, and our main aim was to make some of these practices visible and encourage discussion about them. We identified two themes in these practices that are not immediately evident in the answers of our interviewees. These themes could be developed further. For example, in both cases there seems to be a 'human factor' at play that goes beyond methodology. In 'professionalism' it is the moral loading of the laboratory situation, in 'the production of good data' it is the fickleness of the participants and their possible reactance. This seems a promising direction for further theorising about these themes.

Our study raises the question whether these informal practices make a difference. Most cognitive psychologists thought that they do not have any effects on the outcome of their experiments. Social/organizational psychologists, on the other hand, were more convinced that their informal practices were relevant, although some just did not know. One way to study this question would be to conduct experiments in which an informal aspect of the procedure is itself a variable. One could for example employ two different ways of phrasing or formatting the instructions (without changing their actual content). Although this would double the required sample size and hence might seem unattractive from a pragmatic perspective, this would –in our view– be more than compensated by the knowledge gained.

Another limitation of the current study is that, given its small sample size, it does not allow generalizations about the distribution of informal research practices. Such generalizations were not our aim, but future research may build on our results with a quantitative survey of a larger sample of researchers, using our inventory as a basis for the survey questions (see e.g., Detert & Edmondson, 2011). This would also allow one to test for group differences (e.g., between disciplines or countries).

Conclusion

As mentioned above, the current study harks back to the research on experimenter effects and the ways in which experimenters may inadvertently and unconsciously affect the results of their experiment. Of course, as witnessed by the Many Labs project, research on topics like experimenter effects is all the more relevant in light of the 'replication crisis': If replicability of results is an indicator of their reliability, it is important to identify the possible boundary conditions for attaining certain kinds of effects or results (Derksen & Rietzschel, 2013). The results from our interviews suggest that most of our respondents would agree that differential results *may* be due to small (and unmentioned) differences in a study method, not all of which may be picked up on by a project like Many Labs. The field would benefit from more discussion about and study of these informal practices. The researchers we spoke to were often uncertain whether their informal practices actually make a difference, and unaware of the informal practices of others. However, the interviews acted as a spur for reflection, and despite the initial suspicion and skepticism of some of them, most interviewees seemed to enjoy and appreciate the opportunity to talk about this topic. We hope that the current paper may provoke debate in the discipline as a whole and thus contribute to a climate of openness about research practices, and hence to better research.

Data Accessibility Statement

The data acquired in this study consists of interview recordings and the transcriptions thereof. These are stored on our university's data repository. They will be made accessible on request if questions arise regarding research integrity (e.g. have these interviews really taken place). Because we believe these data can only be interpreted properly with knowledge of the context in which they were produced they will only be shared for reuse with the involvement of the interviewer(s).

Notes

[1] See for instance the debate between Van Bavel et al. (2016a, 2016b) and Inbar (2016).

[2] We planned to do 20 interviews, and approximately 3 per university. Three interviews (20, 21, 22) were conducted as a pilot by MD and ER (together), and 1 pilot and 19 other interviews by JB. We excluded the pilot interview of JB because it was with one of the authors (ER), but included the three pilot interviews of MD and ER because they contained relevant information. We contacted and conducted interviews with researchers from the University of Groningen, University of Amsterdam, VU Amsterdam, Utrecht University, Eindhoven University of Technology, Leiden University and Tilburg University. We choose for these universities because of the psychological research groups and the travel distance.

[3] In terms of informal research practices, this grouping is more relevant than any differences between social and organizational, and cognitive and experimental psychology. We will indicate the first group with the term 'social' and the second group with the term 'cognitive'.

[4] Codes in order of occurrence (from 124 to 2 quotes): Implicit theory, implicit practice, open science, transmission, colleagues, instruction, participants, attitude towards participants, own learning process, good research practices, literature, handbooks, replication, professional, deception, past-present, exclusion, motivation, debriefing, information, performance, own style, checks, standardization, design, scientific system, recruitment, intellectual property, exchange, tacit knowledge, implicit knowledge, approximately, experience, differences social-cognitive psychology, lab environment, natural situation, evident, imagination, intuition, technology, netjes, protocol, assumption, instruction (experimenter), recording, robust, guidelines, participant's experience, naive participants, lay out, annotation, success, time, habits, attitude experimenter, noise, atmosphere, credibility, psychology.

[5] See a.o. Derksen (2001) and Lezaun (2007) for studies of how researchers try get participants to behave naturally.

[6] We thank an anonymous reviewer for raising this question.

Additional File

The additional file for this article can be found as follows:

· **Appendix.** Information about the interviewees. DOI: https://doi.org/10.1525/collabra.221.s1

Competing Interests

The authors have no competing interests to declare.

Author Contributions

· JB: 19 interviews, coding, analysis, writing
· MD: conception and design, 3 interviews, coding, analysis, writing
· ER: conception and design, 3 interviews, writing

References

American Psychological Association. (2010). *Publication manual of the American Psychological Association.* (6th ed.). Washington, DC: American Psychological Association.

Bavel, J. J. V., Mende-Siedlecki, P., Brady, W. J., & **Reinero, D. A.** (2016a). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences, 113*(23), 6454–6459. DOI: https://doi.org/10.1073/pnas.1521897113

Bavel, J. J. V., Mende-Siedlecki, P., Brady, W. J., & **Reinero, D. A.** (2016b). Reply to Inbar: Contextual sensitivity helps explain the reproducibility gap between social and cognitive psychology. *Proceedings of the National Academy of Sciences, 113*(34), E4935–E4936. DOI: https://doi.org/10.1073/pnas.1609700113

Braun, V., & **Clarke, V.** (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. DOI: https://doi.org/10.1191/478088706qp063oa

Buskist, W., & **Johnston, J. M.** (1988). Laboratory Lore and Research Practices in the Experimental Analysis of Human behavior. *The Behavior Analyst, 11*(1), 41–42. DOI: https://doi.org/10.1007/BF03392453

Cohn, S. (2008). Making objective facts from intimate relations: the case of neuroscience and its entanglements with volunteers. *History of the Human Sciences, 21*(4), 86–103. DOI: https://doi.org/10.1177/0952695108095513

Collins, H. M. (1985). *Changing order: replication and induction in scientific practice.* London etc.: SAGE Publications.

Derksen, M. (2001). Discipline, subjectivity and personality: An analysis of the manuals of four psychological tests. *History of the Human Sciences, 14*(1), 25–47. DOI: https://doi.org/10.1177/095269510101400102

Derksen, M., & **Rietzschel, E.** (2013). Surveillance Is Not the Answer, and Replication Is Not a Test: Comment on Kepes and McDaniel, How Trustworthy Is the Scientific Literature in I-O Psychology? *Industrial and Organizational Psychology, 6*(3), 295–298. DOI: https://doi.org/10.1111/iops.12053

Detert, J. R., & **Edmondson, A. C.** (2011). Implicit Voice Theories: Taken-for-Granted Rules of Self-Censorship at Work. *Academy of Management Journal, 54*(3), 461–488. DOI: https://doi.org/10.5465/amj.2011.61967925

Friedman, N. (1967). *The social nature of psychological research: the psychological experiment as a social interaction.* New York.

Gibson, S. (2019). *Arguing, Obeying and Defying: A Rhetorical Perspective on Stanley Milgram's Obedience Experiments.* Cambridge: Cambridge University Press.

Inbar, Y. (2016). Association between contextual dependence and replicability in psychology may be spurious. *Proceedings of the National Academy of Sciences, 113*(34), E4933–E4934. DOI: https://doi.org/10.1073/pnas.1608676113

Kahneman, D. (n.d.). *Kahneman Commentary.* Retrieved from http://www.scribd.com/doc/225285909/Kahneman-Commentary

Kerr, N. L., Ao, X., Hogg, M. A., & **Zhang, J.** (2018). Addressing replicability concerns via adversarial collaboration: Discovering hidden moderators of the minimal intergroup discrimination effect. *Journal of Experimental Social Psychology, 78*, 66–76. DOI: https://doi.org/10.1016/j.jesp.2018.05.001

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Reginald, B., Adams, J., Alper, S., Nosek, B. A., et al. (2018). Many Labs 2: Investigating Variation in Replicability Across Sample and Setting. DOI: https://doi.org/10.31234/osf.io/9654g

Latham, G. P., Erez, M., & **Locke, E. A.** (1988). Resolving scientific disputes by the joint design of crucial experiments by the antagonists: Application to the Erez–Latham dispute regarding participation in goal setting. *Journal of Applied Psychology, 73*(4), 753–772. DOI: https://doi.org/10.1037/0021-9010.73.4.753

Lezaun, J. (2007). A market of opinions: The political epistemology of focus groups. *Sociological Review, 55*, 130–151. DOI: https://doi.org/10.1111/j.1467-954X.2007.00733.x

Peterson, D. (2016). The Baby Factory. Difficult Research Objects, Disciplinary Standards, and the Production of Statistical Significance. *Socius: Sociological Research for a Dynamic World, 2.* DOI: https://doi.org/10.1177/2378023115625071

Srivastava, S. (2014, July 1). Some thoughts on replication and falsifiability: Is this a chance to do better? Retrieved 5 June 2015, from https://hardsci.wordpress.com/2014/07/01/some-thoughts-on-replication-and-falsifiability-is-this-a-chance-to-do-better/

Srivastava, S. (2016, August 18). Lots of us, probably all of us, have 'lab lore' [Micro blog]. Retrieved from https://twitter.com/hardsci/status/766317950059945985

Wilson, A. (2014, May 26). Psychology's real replication problem: our Methods sections. Retrieved 27 May 2014, from http://psychsciencenotes.blogspot.co.uk/2014/05/psychologys-real-replication-problem.html