Can Plans and Situated Actions be Replicated?

John Rooksby
School of Computing Science,
University of Glasgow, UK.
john.rooksby@glasgow.ac.uk

ABSTRACT
This paper discusses a repetition of a study presented in Suchman's book Plans and Situated Actions. There have been complaints about the lack of replication studies in disciplines related to CSCW (particularly Software Engineering and HCI). However, these complaints often become embedded in wider attempts to install a principled scientific method within these disciplines. Plans and Situated Actions was not a scientific text but drew upon naturalistic analysis. This paper shows there is value in recreating Plans and Situated Actions, and argues it would be helpful to recreate other studies. However, such repetition does not and need not constitute a scientific replication. The paper argues that while repetition and reanalysis may improve rigour in computing research, this need not be with a view to making it more scientific.

Author Keywords
Replication; Qualitative Methods

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
This paper addresses an argument often made in Human Computer Interaction (HCI), Software Engineering, and related disciplines: *there ought to be more replication studies* [11, 12, 17, 18, 20, 22, 37, 38, 39]. Replication studies are not unheard of in these disciplines, but according to John [17], they are too often seen as a second-class form of research: researchers do not think their time is best spent conducting or writing up replications, and reviewers do not receive replication studies positively. There are also complaints (e.g. [12, 22]) that authors do not adequately facilitate replications; they do not routinely supply code, data, or more than a cursory overview of their research methods and instruments. In Software Engineering, Gomez et al [11] have proposed several forms of replication study that they believe the discipline would benefit from. Replication studies can, they say: 1) follow the same method as the original study to see if the results can be reproduced; 2) follow a different method to see if the results can be reproduced; and/or 3) reanalyse existing data. In HCI, Wilson et al [37] have produced a broadly similar typology of replication. They argue there can and should be: 1) direct replications, attempting to confirm results of a study; 2) conceptual replications, attempting to replicate findings via alternative studies; 3) 'replicate and extend' studies, which reproduce a study and then build upon results; and 4) applied case studies, that attempt to reproduce results from a study in a real world setting.

The calls for replication studies are often embedded in a broader idea: *HCI and Software Engineering ought to be more scientific*. This idea is usually met with resistance (e.g. [13][35]), largely on the grounds that these disciplines address not technical, but sociotechnical issues. Such resistance is rooted in longstanding doubts about the applicability of scientific methods to social life [1, 16, 23]; it is a mistake to believe scientific methods are universally applicable and are the only valid source of knowledge (a belief that is often described, pejoratively, as scientism). To resist this is not to resist the idea that the quality of research in HCI and Software Engineering can be improved, but to resist the assumption that quality is only achievable through the application of methods from the sciences. Naturalistic and discursive methods offer an alternative of at least equal importance [13, 15, 36]. On these grounds, Greiffenhagen and Reeves [13] are critical of programmatic attempts to introduce replication. They fear such attempts will lead to qualitative approaches being devalued, pointing in particular to ethnography. The issue is not that ethnography is ignored, but that it gets evaluated and potentially transformed according to a positivist or falsificationist agenda with which it does not fit.

This paper revisits a classic book, Suchman's Plans and Situated Actions [35, 36]. The book is critical of cognitive science and presents a naturalistic alternative rooted in conversation analysis and ethnomethodology. The book reports a study in which pairs of participants were videoed using a photocopier. This paper reports a recreation of Suchman's study, arguing it is not possible to replicate Suchman's study (in the scientific sense of the word), but contending that: 1) this does not give reason to dismiss Suchman's findings; and, 2) it is possible and helpful to recreate Suchman's study. The paper concludes that while a scientific-replication agenda can and ought to be resisted from the standpoint of naturalistic research, prescient issues are subsumed within it. If we chose to throw out

CSCW '14, February 15–19, 2014, Baltimore, Maryland, USA.

replication, we should not throw out the possibility of recreating studies.

PLANS AND SITUATED ACTIONS

Plans and Situated Actions is seen by many as a key text on how people interact with technology. The book argues that approaches to HCI that draw from cognitive science (approaches that were dominant in the 1980s when the book was written, what Rogers [29] refers to as "*classical cognitive approaches*") overlook the practical, embodied and organised ways in which people work with and make sense of technology. Suchman argues it is through human action that machines are imbued with sense, and therefore human action and interaction need to be taken into view. To do so we need to look beyond the specific points at which detectable actions (e.g. button presses) occur and at just what it is the users are doing when they operate technology. Suchman argues usability is not a case of building self-explanatory technologies, but technologies that are explainable, comprehensible and amenable within the uses they are put to. She calls for attention to be paid to what she terms the *accountability* of technology.

Plans and Situated Actions was inspired by the problem that Xerox customers were complaining that their photocopiers were difficult to use. Something was wrong with usability design at Xerox. The majority of the book is discursive, focusing on the problems of cognitive science based approaches to HCI and the alternative supplied by naturalistic analysis rooted in ethnomethodology and conversation analysis. The final chapters of the book report a study of how people use a photocopier. This study demonstrates Suchman's arguments, providing not the basis for them but evidence and application. The study does not drive the book as such, but appears at the end, tying together what was written earlier and making explicit the claims she had made. Neither does it use the specific photocopiers the customers were complaining about. Suchman's ideas are not derived from the study, but the study presents a case of them.

For the study, Suchman invited participants to use a photocopier installed at Xerox PARC that had an experimental expert system connected to it. The expert system would lead users through the process of making photocopies. The users would begin by selecting their objective, and the machine would then display text and diagrams to guide them through an appropriate process. The participants were invited to work together in pairs so that they spoke aloud as they were copying. Four pairs were videoed for the study, in sessions that lasted around an hour and a half each. The study showed that the kinds of trouble users were having in making photocopies were not solved by, and in some ways were exacerbated by, the ostensibly intelligent expert system. The users had to make sense of the photocopier and the instructions in ways other than they would have when interacting with a person.

Recreating Plans and Situated Actions

This paper recreates the study reported in Plans and Situated Actions. The study was recreated out of interest in the original text and in order to pursue issues related to the fact the original was a qualitative study done in the laboratory rather than the field. Recreating the study was not straightforward. The specific details of the study are not always clearly presented in the book, and had to be pieced together from various comments and footnotes as well as later works. There are actually three versions of the text: Suchman's thesis [33]; the book (which had minor changes from the thesis) [36]; and a second, revised edition of the book [35]. There is also a transcript of a speech about the study [34]. An edited video that mixes scenes from a film with footage from Suchman's study is also in circulation (although to my knowledge remains unpublished). The study also had to be interpreted in a way appropriate for a new context. In particular, a modern, 'ordinary' photocopier had to be used instead of the experimental system used by Suchman.

THE STUDY

The study was recreated in the computer science department of a university in Scotland. Suchman's book does not summarise the tasks given to the participants, although her transcripts do contain information about them. Most of Suchman's examples are of making a double-sided copy from a bound document, this being a two-step procedure where a single sided copy is made first and then transformed into a double-sided copy. Similar tasks were devised for this study, including one where the participants were instructed by the machine in how to resolve a paper jam. However, as will be discussed, this paper focuses on a simple task involving copying and stapling A4 paper.

The machine used for the study was an Océ 3145 Multifunctional Printer. This machine functions as a photocopier, printer and scanner. For photocopying, it enables the user to place an original directly onto the glass, or to place documents into an automatic document feeder on top. The machine is controlled via a graphical user interface with buttons above and below. Figure 1 shows the interface with the default menu (the basic menu), and the initial instructions displayed when a paper jam is detected. Unlike the machine used in Suchman's study, the Océ 3145 does not lead users through the process of making their photocopies. Only in the event of a paper jam is the user instructed through a process to resolve this. For the most part, the interface enables users to select options as they wish, and press the green start button when ready. Some written instructions are provided on the machine itself, including on the automatic document feeder. Error messages are also displayed on screen in the event that invalid options are selected.

In the book, Suchman states she feared proficient users of the photocopier would mask the phenomena she was looking for and therefore she invited novice users to

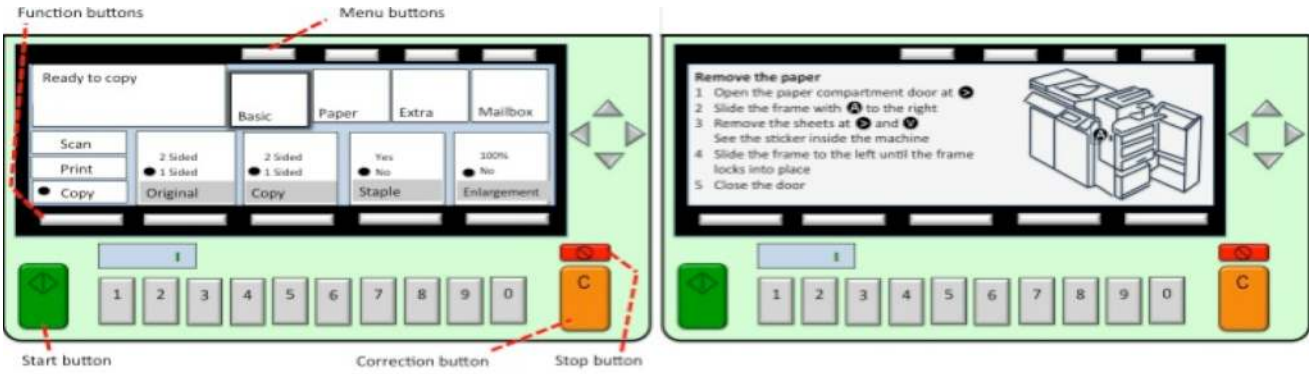

Figure 1: Sketches of the Interface, Showing (1) The "Basic" Menu (2) The Initial Instruction Screen for Paper Jam

participate. She invited people to work in pairs so they would talk aloud as they worked. The new study was run with nine pairs of participants. The study was run five times with postgraduate students and then four times with pairs of academics (at UK-professor, lecturer and post-doctoral level). This paper uses examples from the sessions run with academics. The academics were invited given Suchman's study was revealed to have used academic and research staff associated with Xerox [34]. By coincidence, not design, each of the academic pairs had one person that was experienced with the photocopier used for the study, and one that could be described as a novice.

FINDINGS

This paper will concentrate on how the four pairs of academic staff worked on the first and most simple task they were set: "Make a one-sided copy of Document A with a staple at the top left" (document A being an academic article printed single-sided on A4 paper, in a portrait orientation). The transcripts below detail how the pairs completed this task. The transcripts are presented using a simplified version of the format used by Suchman: (( )) – Double brackets contain an action or comment; ((**bold**)) – Writing in bold denotes an action available to the machine; CAPITALS – capitalisation denotes words said in a loud voice; Heh – denotes laughter.

Pair 1

Pair 1 consisted of a senior lecturer (C) and a postdoctoral researcher (D). They spent around 90 seconds on the task. This pair, as can be seen in the transcript below, engaged in some guesswork but completed the task on their first attempt. Interestingly, after making a successful copy, they stepped back over their actions as if to repeat for their own benefit the successful process. Another interesting aspect, which I will return to later in a discussion of the artificialities of laboratory studies, is that D spoke to the camera twice.

Transcript 1: Pair 1, Task 1

C Are you experienced with this machine?

D I am. Are you?

C Not at all. Perhaps you should start then.

D Alright ((*C hands D the task sheet and document A*)), I'll do the first one then. Right, so this is how I would do it. ((*looks up at camera*)) HELLO ((*looks at machine*)).

C Heh

D ((**places document A in feeder**)) Stick it in the feeder. Copy ((*points at option*)). Original is one sided. Copy is one sided ((*points at option*)). Ah staple. ((*steps back to look at task sheet behind them on table*)) It needs a staple in the top left.

C Do we know what's going to be top left?

D It doesn't ((*shakes head and smiles*)). Yeah, so, we're just gonna see what happens ((**selects staple option**)). We may have to do more than one. I shall attempt to staple. If we can't there's a stapler there ((*points to manual stapler*)). Heh ((**presses start**)) So, we can definitely do a one sided copy.

C And it will put a staple somewhere ((*leans in to watch original coming out into output tray*))

D That's the original coming out ((*C steps back*))

C ((*Copy moves into output tray,* **removed** *by C*)) Genius.

D Aha ((*glances at camera*)) CHECK THAT OUT.

C You put that in ((*points at document feeder*))

D Yeah, that way ((*places document over feeder*))

C Ok, so the staple goes there, yeah ((*aligns copy with original and pinches at staple*)). Ok.

D ((*removes document A from machine*))

Pair 2

Pair 2 consisted of two lecturers (E and F). They completed the task on their second attempt, taking about 160 seconds in total. Their first attempt had a staple at the bottom corner, leading them to reject the copy and try again. Interestingly, after placing document A on the document feeder for the first time, they picked it up again and replayed their actions prior to actually pressing the start button. This was in anticipation of an error – they wanted

to remember their moves so as not to make the same mistake twice. When the error was made, they carefully stepped through what they did, rotating both the original and copy to work out the correct orientation. I will also return later in this paper to the fact that F referred jokingly to the task sheet as a "specification document".

Transcript 2: Pair 2, Task 1

E I've never used a photocopier before like this. So this will be interesting heh heh heh

F Ok

E I didn't even know you could staple things in a photocopier

F Yeah

E But apparently you can

F You can, so eh, ((*leafs through document A*)) this is a one sided document. So I suppose, image side face down ((*points at instruction on document feeder*))

E Right

F And then we need to position it so that if this thing staples it at the end, the staple is over here right? ((*points to lower left corner of document feeder*)

E In the top left, what does that mean? Top left. Right. Ok, err I have to turn on stapler first I guess.

F Ok so, put this in ((**slides document A into document feeder**)), this bit I know

E Right

F And you've got under basic ((*points at menu*)). Original one sided ((*points at menu*)). Stapler yes ((**selects staple option**)).

E Ok

F That's all you need. And then we could just go and do a run ((*points at start button*)) and see if it staples it, on the correct side.

E Ok

F And then work out from what we're getting from which way around we need to turn it. Ok so we've got it ((**removes document from document feeder** and holds up)) so that we're facing and it tilts in towards us ((**places document back into feeder**))

E Ok

F OK ((**presses start**)) with a bit of luck its going to do the right thing. Although I suspect not heh heh

E Heh heh heh

F ((*leans across in anticipation of output*))

E ((**takes copy from output tray** when it emerges))

F Where's the staple

E Err

F Its at the bottom, ok

E Other way round

F So, ((*holds up document A in same orientation as previously, E holds up copy and pinches staple*)) I had this in front of me and was tilting it towards me ((*rotates paper into opposite position and* **places into document feeder**)).

E Yep So it ended up in that corner ((*rotates copy to align with original, and pinches where staple ought to be*))

F Would that be right?

E I don't know heh

F Heh heh ((**presses start**))

E I didn't see the first way you put it, but yeah, ok

F Yeah you can either spend brain cycles or you can waste paper.

E Yeah you could waste paper.

F I'm sure you could do both as well ((*leans over to output tray)*)

E ((**Removes** *output when it emerges and shows to A*)) Dah dah dah

F That complies with the spec. A one sided copy of document A with a staple in the top left.

E Ok ok next one.

Pair 3

Pair 3 consisted of a professor (G) and post-doctoral researcher (H). The pair spent around 180 seconds on the task. The session was somewhat error prone and chaotic. H made several mistakes such as thinking the photocopier had a touchscreen and pressing the wrong buttons. Their first attempt produced a copy with no staple. This was because, after correctly selecting the necessary options, H mistakenly pressed the correction button before pressing the start button, undoing their earlier selections. Their second attempt was successful. Interesting aspects include H's difficulties in understanding the machine, and G's comments to and about the video camera.

Transcript 3: Pair 3, Task 1

G I foolishly consider myself actually quite skilled with photocopiers

H Ok great

G In terms of being able to kind of dig through menus and having a,

H Alright

G Having a, having a silly ability to, try

H To work it?

G Well, to try at least, to kind of, theoretically, ok. make it, so I say we'll

H Document A

G Document A. Make a one sided copy of document A with a staple in the top left. Ok. Document A. Top left. So. Automatic feeder says image side face down, maximum 50 originals to start with, use only

undamaged originals. Looks pretty good, so it's a single sided. Do we have to make it single sided?

H Yeah err it doesn't, one sided copy

G A one sided copy

H Yeah

G So it's a one sided document

H Staple, on the top left yeah

G Image side face down ((**places document in feeder**)) happy with that?

H Yep

G ((*Moves finger over options*)) Now this is where we have to pick the staple, so that's with staple ((**selects staple**))

H Oh I didn't even know it. A4 paper. ((*presses screen where it says paper*))

G Yep

H So, no no these are the menus

H Oh, there's no touchscreen?

G No, there's no touchscreen

H Is this the right button ((**selects original**))

G Yep so its two sided now ((H **selects original**)) that's one sided

H Ok, so we're going, I guess ((**selects paper**))

G Oh

H Because we want A4 paper

G Yeah we want A4 paper. Uhu

H Ok

G Looks good

H ((**Presses correction button**))

G And again probably, oh, no, that's that's clear.

H Oh

G That's ok

Heh ok. ((**selects paper**)) So this is all good.

G Uhuh, just cancel that

H ((**Presses start button**))

G Here we go. So if we've got it set up incorrectly you know, it will pop out with all candles printed on it or something

H ((removes original)) It looks good though

G No that's the original, copies gonna come out there

H ((**removes copy**)) yeah.

G But, it didn't staple it, for some reason.

H Oh it can staple it automatically?

G Yeah it can staple it automatically, so I think I know what happened. So I'm just going to cancel that. ((**presses cancel**, **presses correction**)) Ok so I think it's the staple. ((**selects staple**)) So if you pick this one and then clear it, it wipes all the settings.

H Right, right

G So I think if we do this. ((**presses start**)) It should now work.

H So this is the menu the first row of the settings?

G So these buttons control these menus, and these buttons control these menus

H Right right right right

G And once you've actually got it in the state press copy, and its that one to copy ((*points at start button*))

H Its kind of like

G Yeah I don't quite know. Its just a basic one.

H ((**removes copy**)) Oh yeah, its stapled

G So that is right, its top left. ((reads task sheet)) Make a one sided copy of document A, with a staple on the top left.

H Yeah

G Lets have a look

H That's right

G ((*Turns To camera*)) HELLO, YOU SEE THIS? ((*Turns back*)) This is us using the screen, I think that's what they want. Is that a copy of it?

H Yeah

G Is that the final paper, 969, no I mean the last page. I wanna make sure we've got all the copies.

H Yep

G 969

H Thank you

G Ok, perfect, I'll put the original back into the. Uhuh.

H Oh this is the.

G Original

H No this is the bad one.

G Oh, throw that away. The original's here.

Pair 4

Pair 4 consisted of a lecturer (J) and a postdoctoral researcher (K). They spent about 120 seconds on the task and were successful on their first attempt. This pair were less cooperative. They spoke less and almost fought over the machine. For example, an attempt by K to open the photocopier lid rather than place document A on the feeder was intercepted by J. J held the lid down, and waited for K to infer his preference to use the feeder. This episode differed to the others in that there was no discussion at the outset about respective skills and experience with the copier, and no decision about who would lead.

Transcript 4: Pair 4, Task 1

J Ok, so lets do this.

K So the first one.

J ((*Holds task sheet in hands and reads*)) A one-sided copy of document A ((*puts hand out, and K gives him document A*)). With staple at the top left.

K ((**Opens lid**))

J ((*Places hand on top of lid, restricting further opening*)) Wait wait, wait a second. ((*moves hand over autofeeder*)) So err.

K ((**Closes lid**)) ok

J Soo

K ((*Points to autofeeder*)) face, down

J Yeah. So it can staple them. Erm I think you put it this way ((*places* **document on autofeeder**)). Erm. Face down. And one sided. And with a staple at the top left. So,

K Ah, this document's ((*lifts top sheet of document*)) already one sided

J Yes ((*puts hand over documents, and K moves hers away*)) So the original is one sided ((*points to one sided original option on copier*)). The copy is one sided ((*points to one sided copy option*)). And ((**selects staple**)) we say, staple.

J ((*Pats down document on feeder then* **presses start**))

((*J and K look at screen, then J moves to output tray causing K to step back*))

J ((**Takes output** from machine)) Ok, top left. Right.

K ((*Lifts top page of original*)) ah, ok ((*Takes original from output tray*))

J So, one is done. ((*hands copy to K*)) OK.

K Yeah. This is done.

ANALYSIS: THEMES FROM SUCHMAN

Suchman's analysis covers what she refers to as "*the adept completion of photocopying*", "*responding to the machine*", and "*repair*". This section will demonstrate that even though the current study differs to the one in the original text (i.e. an ordinary photocopier is used, the tasks are somewhat different, and the pairs of participants each had one experienced and one inexperienced person) it is possible to connect transcripts from task one of the current study to Suchman's themes.

The Adept Completion of Photocopying

A key point in Plans and Situated Actions is that plans are resources for actions. So even when someone has a plan or goal, the actions necessary to reach it will need to be worked out (and often reformulated) in ways that cannot be entirely pre-specified. Suchman showed that even where people making photocopies have a goal and are led through a process, reaching that goal is more than blindly following pre-specified steps. The participants had to adeptly make the process work.

Although each pair had a goal (as specified on the task sheet), none of the transcripts above show a plan or process being articulated in advance of the participants' actions. Rather, planning and acting were mutually articulated. Pairs 1, 2 and 3 first established which member was the more experienced, and then the experienced member decided and explained what they should do. The explanation came as a talking-through of the process, starting with the selection of the document feeder and then through the available options.

The first three pairs each anticipated that something might go wrong, in particular that the staple might appear in the wrong place. D, in pair 1, said "this is how I would do it" when explaining what to do (rather than this is how to do it), seemingly orienting to there being alternative ways for completing the task, and probably to limit his accountability for potential failures. Similarly, F in pair 2 began the process with "I suppose" and "this bit I know", and when ready to press start he said "I suppose we could just go and do a run", limiting his accountability for a potential failure. G, in pair 3, was confident in his instructions but as the machine was making the first copy he warned there would be problems "if we've got it set up incorrectly". So, in the first three pairs, one person instructed the other on how to complete the task, but in a way covering more than the specific actions and the button presses required, and in a way that oriented to and limited accountability for potential failure. Two pairs also replayed or re-enacted their actions. F in pair 2, anticipating a potential problem with the orientation of the paper, did not press start when first ready to do so. F placed the document on the feeder and selected 'staple', but then picked the document up again to memorise how he placed it. Pair 1, upon completing the task, replayed and discussed their actions, confirming what it was they did that achieved the successful copy.

In pair 4 there was no initial discussion of competence. K moved to open the photocopier lid, and J physically intervened, signalling but not announcing an alternative course of action. The conflict between these two occurred not because of an absence of a plan, but because the participants had alternative ideas about where to start. The trouble was not resolved by articulating a plan, but simply by signalling or strong-arming a preferred route.

In the examples we have seen that where the task proceeded without error, or where the task proceeded prior to the establishment of an error having occurred, the participants reasoned and narrated through what they were doing, orienting to possible errors, and occasionally replaying their actions. Planning may have been involved, but a comprehensive plan was never formed or relied upon.

Responding to the Machine

Suchman discusses at length the ways in which the users respond to the photocopier, doing this in order to draw out the point that the machine and the people are not interacting on a symmetrical basis. The users are not taking turns with the machine as they do when taking turns in talking to each other. Given the examples are not of users being guided

through a process, there is less to say about this here, only that the users account for and anticipate what the machine is doing as much as they react to it. For example, in pair 1, when the original was taken into the feeder after D had pressed the start button, D commented "so we can definitely do a one sided copy". This described what was happening, but in a way that was open to possible problems with the staple. Next, C leaned in to watch the original coming out, and D remarked "that's the original coming out". This was another narration of what was happening, done in a way to give sense to the output and head off any possible confusion by C. Similarly, in pairs 2 and 4, the person who pressed start stepped across to the copy output tray in anticipation of the copy coming out. F in pair 2 placed his arm across the original when doing so. In pair 3, H picked up the original thinking it was the copy, and G had to comment "no that's the original, copies gonna come out there".

Certain other things could only be established once the machine had run. In particular, the output supplied answers to whether the copy would be correctly stapled. Other reactions to the machine were more surprise ridden, for example pair 3 were surprised when the copy had no staple at all.

The interaction with the machine was asymmetrical. The participants did not take turns with the machine but can be seen to: orient to and anticipate what it will do; narrate, work over, and work out what it is doing; and try to figure out what it has just done. The machine was not self-explaining but was explained – it was in Suchman's terms "*accountable"*.

Repair

Suchman also focuses heavily upon the ways in which people diagnose and repair actual or potential problems and errors. Several errors can be seen being repaired in the transcripts. In pair 3 the screen was mistaken for a touchscreen. This was immediately recognisable to the participants as an error and could be corrected for immediately. A misidentification of the correction button as the start button also appeared, at first, to be immediately repairable. However, the full effect of this error only became apparent later when the output had no staple. At that point, G realised the mistake "I think I know what happened". Here, it can be seen that the sense of prior events and their status as erroneous or correct can be formed through subsequent events. At the point where the unstapled copy was output, the easiest thing to do would simply have been to staple it using the stapler on the side of the machine, but it is not the intention of this paper (or Suchman's book) to consider errors as being deviations from the best or quickest ways to proceed, but to treat these in terms of how these are anticipated, recognised and made sense of by the users.

The most prominent form of error (and repair) in the transcripts is that of not getting the desired result from the output. The participants heavily oriented to the potential for this, even where ultimately the error was not made. It was only the arrival of a document with the staple at the top left that supplied the participants with evidence that their actions were correct. For pair 3, even the arrival of a stapled copy was not evidence enough of correctness, and they went on to compare the page numbers between the original and copy. For the two pairs that did get the staple in the wrong corner, they went on to repeat the process. In terms of the way they operated the machine, their actions were exactly the same on the first and second attempt. The only change was the orientation by which they placed the document in the feeder. For the participants themselves, this was not a repetition but iteration, where producing the copy was done not just with reference to the possibility of a further error, but also with reference to what was done last time.

ANALYSIS: THE LABORATORY SETTING

One thing that is troubling about the CSCW and HCI literature is that Plans and Situated Actions is discussed in several textbooks and papers as if it were an ethnographic field study [32]. For example, Bentley et al [2] refer to Plans and Situated Actions as the prime example of:

"…ethnographic studies [that] have involved a sociologist observing workers in their environment over a period of several months and hence gaining a deep understanding of the actual rather than the formal working practices."

The study Plans and Situated Actions reports is more akin to a laboratory study than a field study. The book differs to cognitive science not through heading out into the workers' environment, but by taking a naturalistic rather than experimental approach to analysis. The book [36] is critical of laboratory methods insofar as such methods introduce controls and predefined coding schemes that "*presuppose the characterization of the phenomenon studied*". As an alternative, the book draws upon "*sort of uncontrolled experimentation*" in which study participants are videoed as they try to achieve particular tasks with a photocopier. The study is certainly not a laboratory *experiment*, and is perhaps better characterised as a simulation. Suchman does not address this laboratory context in the book. In the light of recent arguments that interaction ought to be studied "*in the wild*" (e.g. [30]), it seems timely to ask: how does the laboratory context feature in the participants' work?

Laboratory factors emerged in the study, and are apparent in the transcripts. Perhaps most obviously, one of the members of pair 1, and one member of pair 3 looked and spoke directly to the cameras. In pair 1, C said to D "this is how I would do it" and then looked at a camera and said in a raised voice "HELLO". In pair 3, G showed the correctly stapled copy to the camera and said in a raised voice "HELLO, SEE THIS?" So, not only did the participants address the camera, they did so with a greeting. The greeting appeared as a marker that something was about to happen that (the participant thought) would be of note to the researcher. In pair 3, the greeting preceded the showing of

the copy and the question "SEE THIS?" In pair 1 the greeting seems to demark the beginning of the actual use of the machine. In pair 1, what Francis [9] calls the double-setting of a simulation became particularly apparent; that C was both showing D "how I would do it" but knowingly also demonstrating to the camera how he would do it. The turn to the camera in pair 3 was more of an aside that, once completed, was explained to H as what the researcher wanted (note – they were not actually asked to do this, in the example G is inferring this for himself). Later in pair 1, when they got the correctly stapled copy, F said in a raised voice "CHECK THAT OUT". Similarly to the turn to the camera in pair 3, there is something victorious about this. All of these turns to the camera are done with a special 'camera voice', are accompanied by a look and appear at the beginnings and conclusions of copying rather than in their midst. So, the camera is very obviously not forgotten by the participants in the study. However, in knowing that the camera is there, this just seems to be something to turn to on occasion rather than something that is pervasively obtrusive on the action. The participants also seem to demark when they are engaging with the camera, and spend the rest of the time interacting and working with their partner.

Clearly, the participants are working in a researcher-provoked manner. The participants treated the task almost like a test. In the words of F (pair 2) they were trying to do something that "complies with the spec". That the participants were trying to fulfil a task is obvious at the start of each session where the task is read out. Also, at the end, not only was the production of a successful copy announced but this copy was double checked against the task sheet. The task sheet itself is a physical object. It was not just read out, but handed and moved around, put down, picked up and so on. One of the problems with the task sheet was that it could be mixed up with the other documents, or simply be in the way as the participants do things with the copier. The original documents were also kept organised by the participants during the session. At the end, of each task, the originals were not placed together with the copy but kept separated and treated with the sort of respect often reserved for other people's property. The participants were asked to work together so that they would verbalise their working. This happened, but with a number of side effects. In each transcript the participants not only talked but collaboratively engaged in the activity of copying. Working together required the participants to occupy a physical space at the copier. This meant the person who placed the original into the feeder was not necessarily the person who collected it from the other end, creating a division of labour. The collaborative nature of the tasks also gave them an air of learning or training. Three of the transcripts involve one participant explaining to the other what they would do. A certain level of humour also pervaded the sessions. This humour was part of how the participants managed their interactions with each other, and how they managed their identity on camera.

The participants' actions, however, were not disconnected from ordinary life. The participants can be seen to think about and discuss their experience with the machine, what they ordinarily do, and, to some extent, to think about how using it on this occasion might benefit them for future occasions. Working collaboratively, working with a task sheet, and worrying about what it is someone else wants is also something that plausibly happens in ordinary life. While the laboratory aspects of the task are in some ways a pervasive influence, in most respects the participants orient to the lab context in regular but often fleeting ways. For example, turns to the camera or to the task sheet occurred at the beginning and completion of a task. That the task is researcher-provoked does not mean the participants somehow do photocopying in some strange or new way, but instead makes them orient to what they have been asked and the perceived reasons for being asked as they operate the machine. It does not seem that the participants act in unnatural ways, but in ways that are natural for a laboratory. As long as we don't pretend that a researcher-provoked study is anything other than that, then there does not seem to be any overwhelming reason for saying that interaction is less natural here than in the field.

There are interesting issues here, and room for further debate – not just concerning Suchman's work, but other laboratory simulations (see [31] for another example). In terms of theory, naturalism does not claim there are particular natural environments but that naturalistic work attempts to stay true to the observations [1][31]. This chimes with Suchman's critique of cognitive science, which was a critique of analytical methods and suggests there is no reason from the point of view of interaction analysis to claim that interaction is different in the lab.

DISCUSSION: WAS THIS A REPLICATION?

This work has recreated the study in Plans and Situated Actions. It was not recreated precisely: a modern, ordinary photocopier was used; the tasks were somewhat different; and pairs with one experienced person were used. But it has been argued that the analytical themes from the original text can still be drawn out (which, given the original was designed to throw light on why Xerox customers were having trouble with their machines, should not be entirely surprising). It has also been argued that recreating the study can throw light on concerns about the influence of the laboratory setting in these themes.

But, was this recreation of Plans and Situated Actions a replication? Does this study bolster arguments that there ought to be more replication? I will argue *no*.

What is Replication?

Calls for replication studies are often embedded in programmatic attempts to make computing research more scientific. In 1992 Greenberg and Thimbleby [12]

bemoaned "*the pragmatic difficulty of substantiating experiments through replication*" in HCI, criticizing the field as "*weak science*". Their complaint was not just with the lack of replication, but the lack of studies even worth replicating. HCI as a "*weak science"* did not propose and test *falsifiable* theories. More recently, Wilson et al [38] have stated replication is "*a cornerstone of science ... [yet] we have almost no drive and barely any reason to consider replicating the work of other HCI researchers*." The term "*science*" here could be taken in a general sense indicative of any research (all papers in HCI, CSCW and Software Engineering can be loosely referred to as scientific papers, and all researchers as research scientists) but, in the light of Greenberg and Thimbleby, might equally be taken in a "*strong*", positivist or falsificationist sense: *computing research can be improved if researchers adhere to a scientific method*.

Replication is often said to be a cornerstone of the scientific method. As Popper [28] put it:

"*We do not take even our own observations quite seriously, or accept them as scientific observations, until we have repeated and tested them. Only by such repetitions can we convince ourselves that we are not dealing with a mere isolated coincidence, but with events which, on account of their regularity and reproducibility, are in principle inter-subjectively testable*".

Popper's work is widely read in computing, particularly in Software Engineering where his concept of falsification is foundational to theories of Software Testing (see e.g. [25]). Falsificationism was developed by Popper in the mid 20th century, contradicting the positivist ideas that had dominated to then. The positivists saw science as progressing through the accumulation of empirically observed facts. With falsificationism, Popper contended that nothing can be shown with certainty to be correct or true, but only shown (or not shown) to be incorrect or false. Falsificationism provided a radical critique of positivism but held on to a foundational premise: Popper believed it was the work of philosophy (indeed a moral obligation for philosophy) to articulate the principles for scientists to follow.

There are competing ideas in the philosophy of science, and Popper's work sits among these. Alternative ideas are apparent in the replication literature. Greenberg and Thimbleby's [12] argument was that replications are significant contributions if they falsify a result, whereas Gomez et al [11] and Wilson et al [38, 39] hold a positivistic line in which replications are to confirm (and perhaps extend) prior work. So, if we were to follow a scientific agenda, what is good scientific practice is not a settled issue. It can also be seen that ideas about the scientific method are not necessarily descriptions of how science is practiced but of how science ought to be. Popper's efforts set out, through logic, the principles by which we can learn and know things about the world. For their part, Gomez et al and Wilson et al outline a framework for a more scientifically sound programme of research, one that envisages new practices rather than characterises existing or emerging practices.

New ideas have emerged in the philosophy of science since Popper [5]. In particular, the rise of the sociology of scientific knowledge has led many to question whether there is or can be an ideal method, and to question what value or use defining such an ideal can have for guiding science as it is practiced. Many sociologists have studied scientists in action. Knorr Cetina [19], for example, studied high-energy physics and molecular biology, pointing out that these sciences are practiced in distinct ways. Several sociologists have considered replication in practice (see [3, 6, 7]), finding that replication studies are less common than might be imagined, and that when scientists do try to replicate others' work they do so with reference to their own programmatic interests and usually with alternative equipment and procedures [7]. Therefore, it is doubtful that Poppers' ideals have been fully implemented in science. The successes of science should not be readily taken as evidence for the correctness of Popper's views, and the progress and discoveries made in science should not be readily attributed to replication. As Greiffenhagen and Reeves [13] have previously argued: to embed replication in a broader programme of transforming HCI into a science misunderstands science. It may be possible to contend that Popper's ideas are not invalidated by the sociology of science because sociologists have only studied "*weak science*". But if this is the case, it seems the replication literature in HCI and Software Engineering ought to be working on a much broader vision applicable to all sciences. To make HCI more scientific becomes not a problem of making it more like the sciences but of creating a trans-disciplinary movement from "*weak*" to "*strong*" science.

Replication and Naturalistic Research

Atkinson and Hammersley [1], Lynch [23] and others point out that the social sciences attempt to ensure validity through the specification and strict adherence to defined methods. They explain this contrasts with naturalistic forms of research, which typically seek rigour not through the methods by which data is collected, but through the ways in which the data can be kept true to during the analysis. Claims are often made in ethnography, for example, that the researcher must be immersed over a period of time in a culture. Alternative claims are made in other areas of naturalistic research. Claims are made in ethnomethodology and conversation analysis, for example, that the researcher is dealing in a recognizable social order.

One example that brings to light some of the issues of replication in the humanities is the controversy that surrounded Mead's book Coming of Age in Samoa [24]. Several years after Mead's death, Freedman [10] alleged that her work was misrepresentative and misinformed.

Freedman had interviewed one of Mead's original informants, who told him that Mead was tricked. The informant said she and her contemporaries had not been as sexually active in their youth as Mead had described. However, Freedman's critique has not been upheld in Anthropology. Critics of his work pointed out that even if he had located the original interviewees, they were now old women with a different status in society. Samoan society itself had undergone radical change. Of particular note, the interviewee had converted to Christianity. Was it not more likely that the interviewee would lie to Freedman (an older man on a brief visit) than to Mead (a woman of a similar age to the informants, and who had spent time living with them)? A major aspect of the defensibility of Mead's work was that she kept very detailed notes. These notes are publically available and have enabled others to point out, for example, that Mead was well aware of the kinds of joking, hoaxing and lying that went on in Samoa (see [26]). It is also possible to show with reference to her notes where the interpretive elements of her work come in – her book was not simply a record, but a comparative discussion of two cultures.

The Mead controversy shows that while it is unusual to discuss anthropology in terms of replication, forms of study analogous to the types outlined by Gomez et al [11] and Wilson et al [37, 38, 39] exist therein: 1) field sites are returned to and are discussed with reference to previous studies (Freedman was not the only person to visit Mead's fieldsite, although he was the only one to directly attack her work); 2) ideas and theories of culture are developed and discussed with reference to cases (in Mead's case Samoa and America); 3) data can be systematically recorded, made public, and is used for secondary analysis. Importantly, the controversy also shows that revisiting and reanalysis can feature significantly in a dispute yet not settle it. When studying people, applying the same methods (e.g. interview methods) will not necessarily elicit the same responses. Orans [26] has claimed that Mead's work would have been better if her theories were falsifiable; if the book were of scientific quality then the 'replications' could potentially have come to definite conclusions. But to say this is not to suggest Mead's methods should have been tweaked or tightened but that they should have been radically different. In fact to treat the revisits and reanalyses as replications at all might be said to misunderstand why and how these were done. They were not naïve exercises in replication that were dashed with the realisation that the original was not falsifiable.

Wilson et al [37, 38, 39] say that qualitative work, particularly ethnography, is difficult to replicate. They are right to say there are practical problems to do with time, access and so on in replicating ethnographic work, but the Mead example shows these issues are not insurmountable. The crucial problem is that nothing can be scientifically proved or falsified in a qualitative study. If someone answers questions differently or fieldsites have different features, then what can be said about prior findings? If something is interpreted differently then what does that say about prior interpretations? It seems that arguments and questions can be forwarded about the quality of prior research but that nothing can be proved or disproved by consideration of cause and effect. This does not mean that ethnographers can never be wrong, but that it is perfectly legitimate for different ethnographers to find different things. Quality arises in ethnography not from being able to find provably correct data, but through undertaking research in ways in which appropriate bonds can be formed, through a reflexive consideration of such bonds, and through writing in coherent and (preferably) transparent ways. The situation in anthropology is discursive, but in several respects the Mead controversy is remarkably similar to controversies described by the sociologists of scientific knowledge. For example, Collins' [6,7] studies of replication in physics show that controversies are not settled though a crucial test, but by revisiting, repeating and a lot of arguing.

Replication and Plans and Situated Actions

Ideas from Anthropology have been extensively discussed in CSCW and HCI (and to an extent Software Engineering). One of the key texts to do so is Suchman's Plans and Situated Actions [35, 36]. It should be clear that Suchman's book was not of the same intellectual tradition as Mead's. In fact, Suchman was highly critical of comparative anthropology. She singled out Gladwin's comparative accounts of navigation, criticising them for simplistically portraying western navigation as plan driven and Micronesian navigation as situated. Instead of examining just how navigation was achieved, Suchman found anthropologists to be seeking overly simplistic ways for differentiating one culture from another.

An aspect of ethnography that Suchman was critical of was the use of fieldnotes. Suchman was troubled by these: firstly on the grounds that notes cannot go into sufficient detail, and, secondly, on the grounds that fieldnotes are not adequately verifiable. The alternative Suchman advocated was video analysis. Video could capture just what it was that people did with technology. It could be viewed repeatedly, and by people other than the original observer. In Suchman's advocacy of video is an advocacy, of sorts, for repeatable analysis. With video, observation is not a one-off affair, and is not something done by the lone ethnographer whose notes must then be trusted. Observations become something to be seen by others. Such ideas have come in for debate and criticisms (see [14]) but suffice to say, the methods Suchman advocated offer an alternative to Mead's.

In criticising ethnography, Suchman was not attempting to introduce something more scientific. On the contrary, her book contains a lengthy attack on scientific approaches to the study of human action. Suchman was particularly critical of cognitive science. The methods that cognitive

science relied on, she complained, erased or obscured the very phenomena she believed HCI ought to address. Suchman's work was not anti-scientific; ethnomethodology and conversation analysis, the area drawn upon by Suchman in her book, has approached science (including replication) in a somewhat idiosyncratic but often supportive way. For example Livingston [21] and Bjelic and Lynch [3] have published accounts of their own replications of classic scientific studies. The points they draw out concern the orderly and embodied ways in which experiments are made to work through the know-how of scientists. Scientific experiments do not and cannot report exactly how that experiment was conducted, but rely on certain competencies of the replicator. For example experiments with prisms cannot specify just how it is one positions the prism and their body to find the light in order to see a spectrum of colours [3]. Ethnomethodology does not criticise or offer an alternative to science but argues that human practices are a constituent and indispensible part of the way science is done. This is problematic only for the human sciences: the scientific study of people requires orderly embodied practice but then erases that in favour of an 'ironic' account of what people do. A social science study of how scientists do replication may itself be replicable, but would give an ironic version of what replication actually involves.

So, while the replication literature in HCI gives the impression that ethnographers have not even begun to think about how qualitative work can be scientifically-replicated, the fact is that replication has long been discussed, and activities that resemble those advocated in the replication literature have long been engaged in. It is well known that in naturalistic research that work cannot be replicated and it is widely understood that this does not invalidate the claims made. There are also longstanding debates about appropriate forms of data gathering and to what extent the documentation of a site or practice needs to stand in and of themselves as representative of that site or practice and therefore analysable by others.

CONCLUSION

Can Plans and Situated Actions be replicated? No, at least not in the scientific sense of that term. But this is not to say Suchman's study was not worth repeating. Doing so has:

- Aided understanding - Although the book is often cited as if it were about instruction following (e.g. [4][27]) it has implications beyond this. The study presented in her book *demonstrated* a widely applicable argument. The book is not about the experimental machine used in the study, but how *any* machine is operated: machines are rendered accountable through the work of the operator. Suchman's study was, after all, inspired by problems real customers were having with mass manufactured machines.
- Aided exploration of further issues - Repeating the study has aided exploration of situated action in a laboratory context. Suchman's arguments were not about studying people in the wild, but about the use of naturalistic methods. There are issues for further debate here, but turning the methods used by Suchman to examine how participants orient to the laboratory aspects of a study does not immediately undermine findings made in a (naturalistic) laboratory study.

Plans and Situated Actions is widely cited, but not always in ways that acknowledge the full implications of the book. It seems that the book needs to be more carefully read and discussed. This can likely be said of many other books and studies too. Recreating studies could play a helpful role in this. This should not be confounded with scientific-replication. Repeating studies will not make computing research more scientific, but may help with increasing rigour. Increased rigour, in this sense, does not mean proving or disproving existing results but having more discussion and debate about existing work.

ACKNOWLEDGEMENTS

I would like to thank the study participants, the anonymous reviewers of this paper, and Fan Zhang (for reading the book with me and helping develop the tasks). I began this study while funded by the LSCITS Initiative. I would like to thank Matthew Chalmers and the Populations Programme for allowing me to continue this work at Glasgow.

REFERENCES

1. Atkinson, P. and Hammersley, M. *Ethnography. Principles and Practice* (3rd Edition). Routledge. (2007).
2. Bentley, R. Hughes, J. A. Randall, D. Rodden, T. Sawyer, P. Shapiro, D. and Sommerville, I. Ethnographically-informed systems design for air traffic control. *In Proc CSCW 92*. New York : ACM, (1992). 123-129.
3. Bjelic, D, and Lynch, M. The work of a scientific demonstration: respecifying Newton's and Goethe's theories of color. In G. Watson and R. Seiler (eds.), *Text in Context: Contributions to Ethnomethodology*. London: Sage Publications. (1992) 52-78.
4. Brown, B., Laurier, E. The normal natural troubles of driving with GPS. *In Proc CHI'12* (2012). 1621-1630.
5. Chalmers, A. *What is this Thing Called Science?* Open University Press. (1976).
6. Collins, H. M. The seven sexes: A study in the sociology of a phenomenon, or the replication of experiments in physics. *Sociology*, 9(2) (1975) 205-224.
7. Collins, H. M., *Changing Order: Replication and Induction in Scientific Practice*. Beverley Hills & London: Sage, (1985).
8. Crabtree, A., Rouncefield, M., and Tolmie, P. *Doing Design Ethnography*. Springer. (2012).
9. Francis, D. Game identities and activities: Some ethnomethodological observations. In Crookall D and

Saunders D (eds.) *Communication and simulation. Multilingual Matters*. (1989). 53-68.

10. Freeman, D. *Margaret Mead and Samoa*. Cambridge: Harvard University Press. (1983).
11. Gómez, O., Juristo, N., and Vegas, S. Replication, Reproduction and Re-analysis: Three ways for verifying experimental Findings. *In Proc of RESER International Symposium on Replication in Empirical Software Engineering Research*. (2010).
12. Greenberg, S. and Thimbleby, H. The weak science of human-computer interaction. *In Proc CHI '92*(1992).
13. Greiffenhagen, C., and Reeves, S**.** Is replication important for HCI? In *Workshop on replication in HCI (RepliCHI), SIGCHI Conference on Human Factors in Computing Systems (CHI)*, April (2013).
14. Heath, C., Hindmarsh, J., Luff, P. *Video in Qualitative Research*. Sage. (2010).
15. Hutchins, E. *Cognition in the Wild*. Cambridge MA: MIT Press. (1995).
16. Hutchinson, P., Read, R., and Sharrock, W. *There is No Such Thing as a Social Science: In Defence of Peter Winch*. Ashgate. (2008).
17. John, B. Avoiding "It's Just Replication". In *Workshop on replication in HCI (RepliCHI), SIGCHI Conference on Human Factors in Computing Systems (CHI)*, April (2013).
18. Juristo Juzgado, N., and Vegas, S. The role of non-exact replications in software engineering experiments. *Empirical Software Engineering,* 16(3),(2011), 295-324.
19. Knorr Cetina, K. *Epistemic cultures: how the sciences make knowledge*. Harvard University Press. 1999.
20. Krein, J., Knutson, C., Prechelt, L., and Juristo Juzgado, N. Report from the 2nd international workshop on replication in empirical software engineering research. *ACM SIGSOFT Software Engineering Notes 37(1).* (2012), 27-30.
21. Livingston, E. *Ethnographies of Reason*. Ashgate Publishing, Ltd. (2008).
22. Lung, J., Aranda, J., Easterbrook, S., and Wilson, G. On the difficulty of replicating human subjects studies in software engineering. *In Proc ICSE '08*, (2008), 191-201.
23. Lynch, M. *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. Cambridge: Cambridge University Press. (1993).
24. Mead, M. *Coming of Age in Samoa: A Psychological Study of Primitive Youth for Western Civilisation*. New York: William Morrow and Co. (1928).
25. Myers, G., Sandler C, and Badgett T. *The Art of Software Testing*. John Wiley and Sons (2011).
26. Orans, M. *Not Even Wrong: Margaret Mead, Derek Freeman, and the Samoans*. Chandler & Sharp. (1996).
27. Orlikowski, W. The Duality of Technology. *Organization Science* 3(3) (1992). 398-427.
28. Popper, K. *The Logic of Scientific Discovery*. Hutchinson & Co. (1959).
29. Rogers, Y. *HCI Theory: Classical, Modern, and Contemporary*. Morgan & Claypool Publishers. (2012).
30. Rogers, Y. Interaction Design Gone Wild: Striving for Wild Theory. *Interactions* 18,4. (2011). 58–62.
31. Rooksby, J. Does Professional Work Need to be Studied in a Natural Setting? A Secondary Analysis of a Laboratory Study of Software Developers. In Petre M and van der Hoek (eds.) *Software Designers in Action. A Human Centric Look at Design Work*. Chapman Hall/CRC Press. (2013).
32. Rooksby, J. Wild in the Laboratory: A Discussion of Plans and Situated Actions. *TOCHI*, forthcoming.
33. Suchman L. Plans and Situated Actions. The Problem of Human Machine Communication. *Xerox PARC Technical Report ISL-6.* (1985).
34. Suchman, L. Work Practice and Technology: A Retrospective. In Szymanski, M and Whalen, J. (eds.) *Making Work Visible. Ethnographically Grounded Case Studies of Work Practice*, Cambridge: Cambridge University Press. (2011). 21-33.
35. Suchman, L.. *Human-Machine Reconfigurations. Plans and Situated Actions, 2nd Edition.* Cambridge: Cambridge University Press (2007).
36. Suchman, L.. *Plans and Situated Actions: The Problem of Human Machine Communication*. Cambridge: Cambridge University Press (1987).
37. Wilson, M. L., Resnick, P., Coyle, D., and Chi, E. RepliCHI—The Workshop. In *Extended Abstracts*. *CHI '13* (2013).
38. Wilson, M., Mackay, W., Chi, E., Bernstein, M. and Nichols, J., RepliCHI SIG: from a panel to a new submission venue for replication. *In Ext. Abstracts CHI'12* (2012). 1185-1188.
39. Wilson, M.L., Mackay, W., Chi, E., Bernstein, M., Russell, D. and Thimbleby, H. RepliCHI - CHI should be replicating and validating results more: discuss. *In Ext. Abstracts CHI'11,* (2011). 463-466.