
Copyright © 2015 ACM

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

Content must not be changed in any way or reproduced in any format or medium without the formal permission of the copyright holder(s)

http://eprints.gla.ac.uk/106672/

Deposited on: 26 May 2015
Deep Cover HCI: A Case for Covert Research in HCI

Julie R. Williamson and Daniel Sundén
Glasgow Interactive Systems Group
School of Computing Science
University of Glasgow
Scotland, UK

Julie.Williamson@glasgow.ac.uk
Daniel.Sunden@glasgow.ac.uk

Abstract
The growing popularity of methodologies that turn “to the wild” for real world data creates new ethical issues for the HCI community. For investigations questioning interactions in public or transient spaces, crowd interaction, or natural behaviour, uncontrolled and uninfluenced (by the experimenter) experiences represent the ideal evaluation environment. We argue that covert research can be completed rigorously and ethically to expand our knowledge of ubiquitous technologies. Our approach, which we call Deep Cover HCI, utilises technology-supported observation in public spaces to stage completely undisturbed experiences for evaluation. We complete studies without informed consent and without intervention from an experimenter in order to gain new insights into how people use technology in public settings. We argue there is clear value in this approach, reflect on the ethical issues of such investigations, and describe our ethical guidelines for completing Deep Cover HCI Research.

Author Keywords
Methodologies, Covert Ethnography, Evaluation Techniques, Ethics, Deep Cover.

ACM Classification Keywords
H.5.1 Information Interfaces and Presentation: Multimedia Systems (Evaluation Methodology).
Introduction

In a time when many research questions lead us to evaluate "in the wild" [9], it seems like the next logical step to evaluate completely undisturbed interaction in the wild. Evaluations and interventions without any interference or visible presence from an experimenter could give us a more realistic view of how our technologies and interfaces are used in practice. Moreover, staging the experiment to remove the experimental setting could make this effect even stronger: the participants may not even realise it is an experiment at all. This would give the most realistic setting for evaluating interaction in the wild, where we can utilise not just "ecological validity" but actual ecological reality.

At this point, the obvious question arises about research ethics. What about informed consent? What about data anonymity? What about research ethics? However, there are clear guidelines for completing such research if we look towards other disciplines. The sociological authorities on ethics have a long tradition of handling situations where informed consent may not be practical or would disrupt the phenomenon in question. Although covert research has long been a debated area of sociology, there are clear motivations for this technique and detailed ethical guidelines.

The Human Computer Interaction (HCI) community has arguably grown out of borrowing and extending (for better or for worse) methodologies and theories from other disciplines. Why not the tradition of covert methods, applied to fit the needs, values, and ethics of this community? Covert research is already being used in certain areas of HCI. For example, seminal papers in the area of public displays, such as CityWall [25] and StrikeAPose [29] could not realistically have collected informed consent from every passer-by that happened to walk near their displays. Interestingly, neither of these papers specifically address the research ethics that played a key role in these seminal evaluations.

We argue there is clear value in completing covert research because these studies are able to reveal new understanding about interaction not possible using other techniques. We also believe there are clear ethical guidelines that must be followed to ensure covert research is done responsibly and rigorously.

Deep Cover HCI is our approach to completing covert research into users’ naturalistic responses to interactive technologies. This approach revolves around intervention-based evaluation, where we stage our research around an installation or technology that has been purposefully placed as opposed to evaluating existing technology use. Because of our initial intervention, we feel this method goes beyond basic covert research to a form of deep cover. After our initial intervention, we aim to maintain cover by avoiding all further intervention or disturbance to the experimental setting.

This alt.chi submission argues for the value of covert research methods in HCI and presents a discussion on the ethics of completing covert research. this approach utilises technology-supported observation, staging in public or semi-public spaces where passers-by would reasonably expect to be observed, and no intervention whatsoever from the experimenter after installation. The main contributions of this paper are:

- Reflection on the history of covert ethnography, covert participant observation and the value of covert research in HCI
Discussion of the core ethical issues and appropriate guidelines for covert research
Presentation of our approach to Deep Cover HCI, with guidelines for practice and reflection on limitations

Participant Observation, Covert Ethnography, and Outsider Roles
Using ethnographic techniques to study familiar settings “close to home” was arguably driven by the urban sociologists of the Chicago School at the turn of the twentieth century [24]. Ethnographic techniques traditionally used primarily by anthropologists in remote locations now became the tools of sociologists seeking to understand the less exotic but equally complex society of urban Chicago. Since its inception, urban sociology has grappled with questions of ethics. One of the early influences of urban sociology was Jane Addams’s participatory work with the residents of Hull House. Her work clearly blurred the lines between social action and research. For example, Addams used the analysis of data collected by the residents of Hull House to advocate for social causes across the city [1].

Using even “traditional” ethnographic techniques requires careful consideration of communicating research motives, maintaining anonymity, and presenting a rigorous and honest portrayal of the subject. These issues can become even more complicated when completing covert research. The balance between the value gained versus ethical issues of such work has been passionately debated since the 1950s [7]. Covert observation, especially covert participant observation has been used in a variety of contexts to understand everything from a closed group of religious extremists [14], a community of door “bouncers” [8], and the mysterious world of sororities [27]. These studies vary in their rigour and sensationalism, the closed nature of their subjects, and the value gained from the use of covert methods as compared to other techniques. But the insights gained provide a compelling example of contexts where covert observation can generate unique understandings.

Bulmer describes the roles of social scientists with respect to their relationship to the subjects (insider or outsider role) and their choice of method (overt or covert) [7]. Of particular interest in this work is the comparison of outsider roles and the resulting experience created by using covert versus overt observation methods in this context. As an outsider, William F. Whyte used overt methods in his investigations of Italian slums [30]. However, William H. Whyte’s use of time-lapse cameras to study how New Yorkers used public spaces is an excellent example of covert research from an outsider’s perceptive [31]. Deep Cover HCI utilises a covert outsider role, positioning the researcher as an outsider as opposed to a participant or insider.

The difference between outsider and insider roles has significant practical and ethical implications. Lofland et al present a discussion of the issues of completing covert observation in public, quasi-public, private and quasi-private settings [22]. For example, there are different ethical consideration when observing passers-by on the street as an outside and observing the inner workings of a closed group as an insider. Lofland et al discuss issues that specifically revolve around participant observation and insider roles, noting that the close scrutiny needed for serious investigation is likely to cause suspicion or break cover in many settings [22].
Appropriation of Research Methods in HCI

As HCI has grown out of computing science, sociology, psychology, design, and many others, the community has drawn theory, methodologies and practice from a variety of sources. The appropriation (and sometimes misappropriation) of core concepts from other disciplines has been the subject of substantial reflection and critique in this community.

There is significant criticism of the way ethnography is used (and arguably at times misused) in the HCI community (for example [10] and [11]). Dourish discusses that reducing ethnography to another “tool” in the requirements toolbox not only marginalises the underlying theory, it fails to capture the serious conceptual contributions that ethnographic enquiry can make as a method in its own right. Another issue with using ethnography simply as a means to justify “implications for design” is that the analytic process is obscured [11]. Major issues for both producers and consumers of ethnographic work in HCI are perceived lack of rigour, lack of detail/understanding of the analytic process, and misunderstanding about how such results might apply to new contexts. There is a clear lack of reflection on our incorporation of ethnography in HCI [10].

The HCI community has a tendency to adopt new research methods in such a way that important theoretical aspects of the method are lost and/or the method is minimised to a discrete step in a predictable and well-ordered process. For example, HCI researchers have consistently misused parametric statistics when Likert-type scales are plugged in to experimental designs as a convenient tool [19]. The cultural probe [16] has been the subject of much discussion over it’s varied uses [5]. These same attitudes often lead to the positioning of “lab studies” as automatic precursors to “in the wild” studies. However, this fails to recognise the distinct results and goals that can be achieved through these different techniques. In the wild evaluations do not necessarily follow from lab studies in a scientifically important or interesting way. We argue that “in the wild” studies can be a standalone technique that provides valid and useful data.

Ethics of In the Wild HCI

Reflecting on general ethical practice in HCI is not a new theme, and one can find extensive ethical guidelines for lab-based research from a variety of ethical authorities, for example from psychology [2][6], with a strong tradition in lab evaluation.

However, the growing use of “in the wild” HCI leads to interesting changes in the ethics discussion as the setting for research becomes less controlled and more varied. For example, evaluation techniques that involve large-scale studies deployed through app stores bring up new issues about informed consent, data anonymity, and participant expectations [23]. Evaluation settings such as public art events blur the boundary between public engagement and research. In the evaluation of Tweetris, researchers grappled with guidelines that allow for observation without an ethics review when the intervention is not invasion or interactive [26]. To deal with this, researchers sought approval for observations without consent but gathered traditional consent for collecting questionnaire data from participants.

An area that naturally lends itself to in the wild evaluations is public displays. Many evaluations of public displays look at “in the wild” interventions but ethics are often not addressed explicitly as part of the results.
Numerous studies have involved participants without explicit consent (for example [25][29][20]) but ethics discussions are noticeably absent. Langheinrich et al explore procedural issues for ethics without specifically tackling ethical practices or guidelines [21]. In a field where there is a wealth of understanding to be gained from covert research, it is clear that it is time for the community discuss these issues and establish rigorous an ethical approach to “in the wild” evaluations.

**Deep Cover HCI**

We argue that there is clear value in completing covert and “deep cover” research in HCI, and that there are ethical ways to complete this research in a variety of settings. For many “in the wild” evaluations, gathering informed consent from every possible participant is impractical and disruptive. For example, when researchers stage evaluations as part of public events (such as [26] [28]), signing consent forms and collecting qualitative data through questionnaires would be heavily disruptive to the experience as well as the natural behaviour of the participants, possibly skewing any data collected. Evaluations concerned with the usability of prompts without guidance [29], the attractiveness of displays [20], and “walk up” experience and appropriation [25] would be difficult to study meaningfully in a lab. Additionally, data collected in a lab would be influenced in unknown ways as a result of overt observation. There are clear scenarios where covert research and evaluation in the wild will provide the most useful and valid data.

Deep Cover HCI is an intervention-based approach to evaluating technology in public spaces, where the technology in question is staged such that passers-by may not be aware they are part of an experimental setting.

The initial intervention puts the researcher into a special position, where the researcher is aware of the intervention and its purpose but passers-by will not be aware of the research. This changes the approach from basic covert research, where a researcher observes peoples’ existing behaviours, to a form a deep cover, where a researcher is actively influencing the setting in their role as a covert outsider. Maintaining cover in this context means maintaining the secrecy of experimental purpose of the installation from passers-by.

The key components of Deep Cover HCI consist of:

- Blurring the lines between experimental settings and real world settings through evaluation staging
- No experimenter intervention or visible presence after initial intervention
- Analysis based on multiple streams of observable data only
- No explicit consent gathered from participants at any point during the evaluation

**Experimental Staging**

A key component of Deep Cover HCI is staging experiments such that participants are not aware of the evaluation or the manipulation of variables. This will require thoughtful reflection on research questions and goals as well as more practical issues such as site access, context, and potential users. For example, the technology must be tailored to the setting or vice versa such that the staging is appropriate and does not evoke suspicion by its presence. Additionally, deployment hardware must be sophisticated enough to run without an experimenter present and support an experience that can be completed without guidance or training.
Not all interfaces or technologies are appropriate or make sense to be deployed in such a setting.

It is important to recognize the limited settings where staging a covert experiment is ethical. In general, the only ethical settings for completing this work are public and quasi-public settings. Additionally, these settings must be places where participants would reasonably expect to be observed at any given time.

**Non-Intervention**
In order to maintain cover for the duration of an evaluation, the experimenter must not intervene unless absolutely necessary for health and safety. The presence of an experimenter not only disrupts the staging of the evaluation, it has significant and unknown effects on the observational data collected.

As part of Deep Cover HCI, we do not feel that collecting qualitative data from passers-by is appropriate or useful. Firstly, this would require an experimenter be present in the deployment space, which may deter potential users or have other unintended effects on observed behaviours. Secondly, we question the value such data brings, especially given the cost of “breaking cover” in the context of the experiment. If such data were to be collected covertly, we question the ethics of such an approach. It would be difficult to complete such covert questioning successfully without creating suspicion or breaking cover. Limiting data collection to observable behaviour increases the reliability of collected data and supports ethical data collection guidelines.

**Naturalistic Observation**
By using technology-supported observation techniques, we can collect data traditionally gathered using naturalistic observation at scale. This data can be collected as a constant stream of input from a variety of sources to support both qualitative and quantitative analysis. Based on the ethical guidelines reviewed in detail later in this paper, limiting data collection to observable behaviour is a reasonable and ethical approach to covert data collection.

A key data source for Deep Cover HCI is behavioural maps generated from video data [32]. Using the data generated from automatic pedestrian tracking, researchers can analyse factors such as walking speed, direction, and time spent during interaction. Behavioural maps also visualize flows of traffic, areas where passers-by crowd, and how passers-by use the space. Behavioural maps support analysis of both interacting and non-interacting users and present data in a completely anonymised format [32]. Additionally, video segments can be used for detailed analysis using a variety of approaches.

Detailed on-device logging can give a view into how users interacted with the technology in question, and can be tailored to the specific research questions and hardware being used. Such logs can provide a detailed portrayal of on-device interaction while maintaining anonymity of users.

Bringing together these data sources, researchers can complete rigorous qualitative and quantitative analysis. To ensure ethical use of the data, results must be visualised and shared such that users can maintain anonymity. No personally identifying images or results should ever be presented.
Unwitting Participants
By following the core principles of Deep Cover HCI, participants should not be aware they are participating in an experiment and thus exhibit uninfluenced responses to and interactions with the technology in question. If the guidelines and restrictions are followed, there should not be any need to obtain informed consent from participants at any point during the evaluation. Researchers must ensure that there is no risk of harm or other negative impacts to these participants.

Ethics
A key ethical issue of completing deep cover HCI research is the lack of informed consent. The ACM Code of Ethics and Professional Conduct describes our “moral imperatives” but consent does not feature except in the context of respecting the privacy of others [4]. The Institute for Electrical and Electronics Engineers (IEEE) Code of Ethics is even terser on the concept of consent [17]. These codes of ethics do not tackle the detailed ethical needs of research practice in HCI. However, there are guidelines for completing such research from a multitude of ethical authorities in the social sciences and humanities. These authorities vary in their attitudes towards covert research, but all agree there are times when such research is necessary and that special precautions must be taken.

The Economic and Social Research Council (ESRC), the main source of social science research funding in the United Kingdom, provides a detailed rationale for contexts where covert research is necessary. The ESRC Framework for Research Ethics states “Informed consent may be impracticable or meaningless in some research, such as research on crowd behaviour, or may be contrary to the research design, as is sometimes the case in psychological experiments where consent would compromise the objective of the research... Covert research may be undertaken when it may provide unique forms of evidence or where overt observation might alter the phenomenon being studied.” [12] The International Sociological Association (ISA) gives a short guideline on covert research, stating that “Covert research should be avoided in principle, unless it is the only method by which information can be gathered, and/or when access to the usual sources of information is obstructed by those in power.” [18] These are the primary reasons for completing covert research, and must be considered before choosing this method.

The American Sociological Society (ASA) gives a more detailed description of settings or contexts where covert research may be appropriate. For example, the ASA states “Sociologists may conduct research in public places or use publicly-available information about individuals (e.g., naturalistic observations in public places, analysis of public records, or archival research) without obtaining consent.” [3] The use of naturalistic observations in public spaces can be completed without consent as a straightforward aspect of social research, where the setting is a key factor. Completing observational research in public settings where people may expect to be observed does not violate privacy. However, concepts of public/private need to be discussed, as we’ll see in the guidelines from the European Commission.

The European Commission (EC), the main funding body for European research, recently completed its Framework Programme 7. This programme generated a large amount of documentation on ethics for research in the social sciences and humanities (which at the time of writing the current programme Horizon 2020 is still yet
to establish). The Guidance Note for Researchers and Evaluators of Social Sciences and Humanities Research states “For example in ‘covert research,’ researchers should take into account the meanings of public and private in the contexts they are studying. Covert observation should only proceed if researchers can demonstrate clear benefits of the research, when no other research approach seems possible and when it is reasonably certain that no one will be harmed or suffer as a result of the observation.” [13] This is the most complete guidance in that it brings together the appropriateness of covert research with respect to setting. These guidelines also state “Another area of ethical concern pertains to the observational research that is central to much sociopsychological research. Observational approaches can vary (focused, participant, invasive/intrusive, visible, covert/overt; recorded rigorously using audio/visual methods or hand written notes compiled after the event). Researchers should ask themselves several questions concerning the research setting (e.g., is it public or private?), the behaviour under scrutiny (in a public or private setting), the way data is collected (recorded or not), and whether or not the protection of participants is ensured.” When completing covert research, we must ensure the anonymisation of data. We aim to achieve this in Deep Cover HCI by limiting data collection to observable data and presenting data in visualisations to maintain anonymity.

In reviewing each of these guidelines, we highlight three key questions to extend basic ethical guidelines that must be addressed to determine if covert research is appropriate:

- Is covert research the only way this data could be collected? For example, is consent impractical or would consent disrupt the phenomenon being observed?
- Is the setting one where people might reasonably expect to be observed? If not, then covert research may have serious ethical issues.
- What kind of data will be collected, and will the research maintain the anonymity of those generating the data?

**Conclusion**

Deep Cover HCI is an ethical approach to covert research that supports rigorous quantitative and qualitative analysis. In the paper, we present the approach and describe the kind of data that can be collected such as pedestrian tracking data and on-screen logging. However, completing analysis rigorously using this data will vary depending on the technology used and the research questions. What we consider “rigorous” in HCI is still an open issue in the community and one that warrants debate.

One of the most debated issues in Deep Cover HCI is the purposeful exclusion of qualitative data collection in order to maintain cover during evaluation. At this point, it is difficult to understand what effects the presence of an experimenter has on the observational data collected. For example, does the behaviour of the experimenter deter others from approaching the display? Does the data collected in this way give an unbiased view into user opinions? Until these questions can be answered, we would argue that collecting qualitative data at a deployment site creates unknown bias in observational data.
Because of the significant effort Deep Cover HCI makes to anonymise data, making data open for review and secondary analysis should be straightforward. Making data openly available is becoming a priority for many research councils, for example the EC Horizon2020 programme recently started its Open Data Pilot. Open data is also important because it adds transparency to analysis techniques and allows for critique of data practice and analysis techniques. Making data publicly available also brings up the question of who owns the data that is generated through observational studies. For example, if a user becomes aware that they generated data in a publicly available data set, do they have a right to ask for it to be removed? Would this even make sense practically or ethically from a researchers’ perspective? Reflection on ownership will form a critical part of any approach to open data.

Limitations
There are clear constraints and limitations of using the Deep Cover HCI approach. Firstly, the settings in which covert research is ethical are limited to public or quasi-public settings. For questions involving private settings, covert research may not be appropriate. Secondly, the type of data that can be collected is limited. Maintaining cover means not approaching passers-by for information. Limiting data collection to observable data also ensure data collection remains ethical, maintaining privacy and anonymity of those generating the data. Thirdly, prototypes for evaluation must be sophisticated and ready for “in the wild” evaluation. Technologies or prototypes that cannot be left unattended are not suitable for Deep Cover HCI.

Acknowledgements
We would like to thank all of the unwitting participants that have fed our curiosity and enabled us to do the work that intrigues us and expands our understanding of technology in public spaces. This work was supported by the EPSRC SIPS project (publicinteraction.co.uk).

References


