# Ethics Issues in Custom Software and App-Based Research Experiments and Data Collection

## Guidance for Researchers and Ethics Reviewers

### Nicolas Gold, Department of Computer Science, UCL

#### Version History

Version 0.1: (22/6/2020) First draft to CS Ethics Committee for review.
Version 0.2: (3/7/2020) Second draft incorporating CS Ethics Committee feedback/suggestions. Circulated to CS Ethics Committee and UCL DPO for comment.
Version 1.0: (16/7/2020) Third and final draft released to CS Ethics Committee and UCL REC for formal consideration incorporating comments received to this point.
Version 1.1: (16/7/2020) Minor typographical amendment. Page numbers added.
Version 1.2: (23/7/2020) Updated to incorporate recommendations from UCL Data Protection team.

This document provides guidance regarding the research ethics issues raised (in general) by the use of custom software and/or app-based data collection. It was produced at the request of the Chair of the CS Ethics Committee. Definitions of terms used may be found in Appendix A.

##### General Issues

* Custom software (in the form of an app or other data collection platform) embodies an approach, workflow, or method to experimentation and/or data collection (e.g. conditional availability of questions depending on prior responses, task-based collection, sensor readings etc). The data collection may be explicit (like a questionnaire) or implicit (e.g. measuring something about a person or their behaviour, or in response to a task). As such, the research method to be used may be opaque to the reviewer, and essentially unbounded in scope and capability. **Ethics documentation thus needs to be clear about what the software is doing and how this has been assured (owing to the inherently buggy nature of software).**
* The quality of the software is important for various reasons: UCL is represented by it, there may be data protection issues if it is not well-engineered, it must engender trust in the users that it is only doing what is expected, it must operate safely and securely (including in any interaction it may have with external hardware like headphones where software errors could cause physical harm), it should not cause significant demands on a participant’s device itself (e.g. battery life, network access) or on the participant’s financial situation (e.g. using paid mobile data allocation) and it should not compromise user interface measures like screen locking (or if it does, this should be justified and consented to clearly). **Ethics documentation should contain details of the design and quality assurance processes that have been applied to meet these criteria (e.g. formal code review from colleagues or UCL Research IT Services or similar).**
* UCL branding and brand management may be required for software release. **Ethics documentation should contain details of how the software will be approved for release by UCL and be officially recognised as UCL software.**
* There must be **clarity about withdrawal from a study and what is involved in doing so**. Uninstall instructions should be provided for participants who wish to withdraw or who have completed a study, statements about the disposition, location, and destruction of data must also be clear (e.g. is data on the device destroyed but that residing at UCL not destroyed?). **Information must also be provided about what will happen should the participant choose not to uninstall the app**.
* There must be a **clear plan** in the ethics documentation for the **provision of technical support** to potential participants (that does not rely on them retaining access to the device in question), **and contingency plans** should things go wrong (e.g. we have heard of a study where a research app was installed and a user’s accounts were subsequently compromised – the user blamed the research app and it was only the serendipitous presence of a technical expert that diffused the situation).

##### Mobile App Issues

* Apps running on a **participant’s own device pose additional challenges to privacy** (apps running on UCL devices are not inherently different to those running on a UCL server or UCL local computer in this respect aside from data security if the device is being used away from UCL).
* Apps to be run on the participant device need to be installed on their device, either through an app store, or through side-loading. **App store release is the safest approach** since it ensures an additional quality control layer (the app checking team/process at Apple or Google). There may be a cost involved in this but UCL should have an institutional-level account that can be used: information on where this held is actively being sought. **App-store release is therefore the strongly recommended approach.**
* Apps running on participants’ devices should **comply with platform controls and restrictions** on what is permitted to apps, and in distribution arrangements. App-store release will help to ensure this. Circumventing such restrictions by using loopholes (e.g. as described here: <http://www.experiencesampler.com/materials/>) in things like app testing programmes is unlikely to be considered ethically sound unless specific agreements are entered into with the companies concerned.
* **Side-loading** will require a participant to permit installation of apps from untrusted sources, thereby **potentially lowering the security of their device**. Since the researcher has no knowledge of the expertise of the participant, no knowledge of whether the participant understands the implications and risks of doing this, or if they know how to reverse the process, ethical controls are difficult to apply. **It is recommended that this approach is avoided and is questioned by reviewers should it be proposed. Side-loading release may not comply with UCL branding requirements** (these are under investigation).

##### Data Protection and Data-Related Ethics Issues

* Apps running on **participant devices increase the complexity of data protection issues**.
* The usual data protection considerations around anonymous and personally-identifiable data apply to participant-device apps as they would to questionnaire and survey studies (e.g. presence of free text boxes), although additional consideration will be needed as apps can collect device data that might be identifying (e.g. location, IP address, mobile phone number, email, contacts etc). The accessible data may include the personally identifying data of people other than the device owner/user and these other people would not have had the opportunity to exercise their autonomy in research participation through a consent process.
* It is **unlikely** that any app-based project running on a participant device can be considered **wholly anonymous** owing to the potential access an app has to on-device data, and if released through an app-store, information that UCL may receive about the recipient’s account. It is therefore likely in almost all cases that projects will need to be registered with the Data Protection Office at UCL.
* Participants may need to **give the app permissions to operate**. Specific permissions (e.g. network access) may be grouped together by the operating system (e.g. combining Bluetooth access with network access and mobile data into a single permission group request covering all of them). The granularity of such ‘permission groups’ may not match the granularity of the data the researcher wishes to access (e.g. they may only need Bluetooth but cannot avoid also asking for permission for network and mobile data), meaning that the **user may have to grant more access to an app than is technically required** for the research data that is to be collected.
* **The participant information sheet needs to be very clear** about what is and is not to be collected, and **all technical steps should be taken and verified to prevent the collection of data that is not required**. Where collecting unwanted data is unavoidable, protocols must be in place to remove it from the data prior to analysis. Researchers must also design for the fact that the location of a particular permission within a permission group may change between versions of an operating system[[1]](#footnote-1) and must therefore consider how their app will operate in this respect across multiple operating system versions that may be in use (or restrict their app to specific versions of the operating system) to ensure that participants are not misled. **The interaction of platform version, permissions, app capability, and participant information may be complex** and thought will need to be given as to how this information can be presented clearly and appropriately.
* **Data** collected by a UCL-provided app **on a participant device is under UCL’s responsibility as a Data Controller from the time of collection** **on that device** and while it remains on that device or otherwise held by UCL (this has been verified with the UCL DPO). Consequently, researchers must ensure that information about the nature of data security is provided to participants in the information sheet, that agreement is consented to, and **must take technical steps to secure the data on the device** as part of the app (e.g. through strong encryption). Participants must be able to exercise their GDPR rights (with due regard to the legal basis for processing and the consented ethical basis for withdrawal), and researchers must ensure that should a device be stolen or lost, no-one other than the research team can retrieve the data (e.g. consider using asymmetric encryption so that data is encrypted on-device with a public key but cannot be decrypted except onsite at UCL using the corresponding private key). Researchers should take advice on the security aspects where necessary.
* Data retrieval from a UCL-provided app on a participant device should be to an **appropriately-secured UCL end point and using appropriate encryption in transit and at rest**. Statements to this effect should be in the participant information sheet and appropriate auditing should take place in liaison with technical staff to ensure that opening such an endpoint does not place UCL’s network at risk of compromise. Researchers should **not** use non UCL-contracted cloud services (i.e. do not do what is suggested here wrt Google: <http://www.experiencesampler.com/materials/>).
* App developers should **read and apply the ICO’s guidance on app development**: <https://ico.org.uk/media/for-organisations/documents/1596/privacy-in-mobile-apps-dp-guidance.pdf>
* All projects should **complete a DPIA** to ensure that developers have considered and documented their legal basis for data processing (including special category data), and have taken any action required to comply with Article 5 data protection principles.
* Developers should **consult UCL’s guidance on the production of privacy notices** (<https://www.ucl.ac.uk/data-protection/guidance-staff-students-and-researchers/practical-data-protection-guidance-notices/guidance-writing>) and cross-refer to the UCL Research Participant notice where possible (<https://www.ucl.ac.uk/legal-services/privacy/ucl-general-research-participant-privacy-notice>).
* **Proposals involving apps to be run on participant devices should be scrutinised for the level of intended data access (e.g. the participant’s data), unintended data access (e.g. data about their contacts) and associated access controls. Reviewers should scrutinise proposals for evidence that these issues have been considered in the design of the research protocol, its implementation in software, the quality control applied, and the way in which participants receive and give consent to the study information (e.g. inside an app and/or through a standard consent form and so forth).**

##### Appendix A: Definitions

* **Custom Software**: software written by, or adapted by, a researcher for the purpose of administering an experiment, collecting data, or that is otherwise involved in the carrying out of research. Sometimes termed an “App” or “Application” and usually run on a laptop or desktop PC, or server. Configurable off the shelf packages or platforms (e.g. Qualtrics, REDCap) that do not require compilation, adaptation of their code and/or building of new software do not fall within this definition (although may still require ethics review for suitability).
* **Mobile App**: a form of software designed to run on a mobile device such as a phone or tablet on operating systems like iOS and Android.
* **App Store**: a distribution method for mobile and non-mobile apps e.g. Microsoft Store, Apple Store, Google Play. Usually managed by large companies and with some degree of quality control and provenance checking applied to the software and the producer of the software (although this varies). Usually regarded by users as a trusted source of software for devices.
* **Side Loading**:A means by which an app can be installed on a mobile device without using an app store. This usually requires the device user to authorise ‘unofficial’ sources for applications and may require a greater degree of technical expertise to do so. The security of a device may be compromised to some extent by this action.
* **Permissions**:Apps often require access to hardware and/or other information about, or on, a device (e.g. contacts, network use, camera use, location etc). These are usually requested by the app in the form of permissions at the time of installation.
1. <https://developer.android.com/guide/topics/permissions/overview> [↑](#footnote-ref-1)