

**Guidelines for the Use of Deception and Debriefing in Research**  
Committee on Human Subjects in Research  
University of Toronto

These guidelines are intended to help researchers to think through ethics issues and follow good practices regarding deception and debriefing of research participants. They should be read as a complement to existing guidelines in the Tri-council policy statement: Ethical conduct for research involving humans (TCPS)—particularly Article 2.1c), and several paragraphs of related commentary, available at:

<http://www.pre.ethics.gc.ca/english/policystatement/section2.cfm#2A>

This document was prepared by Dr. Dean Sharpe, Research Ethics Officer for Social Sciences and Humanities, in consultation with Professor Patricia Pliner, Department of Psychology; it was approved by the Committee on Human Subjects in Research on April 16, 2008.

#### Guiding principles and relevant background

The key over-arching principle in research ethics as it is discussed in the TCPS is respect for human dignity, of which a key corollary is respect for free and informed consent. The general force of this principle is that individuals should be in a position to make a genuinely free and informed choice as to whether or not to participate in a research study. The purpose of this document is to bring out some considerations that can be relevant in applying this general principle.

To begin, it bears mentioning that there is a non-trivial sense in which many, if not most, research projects might reasonably be argued to fall short of telling participants “everything”—given, for instance, the level of expert knowledge involved in conducting the research. While there is no doubt an element of truth to this, it should not be taken as license to under-inform; rather, it should be taken as underscoring the need to aim for a high standard in providing all research participants with opportunities for debriefing, offering results summaries, and even offering courtesy copies of final reports if participants understand themselves to be stakeholders in the research. Such a positive approach toward openness and knowledge transfer is presumably in the interests of both participants and researchers alike, and may rightfully be understood to be part of the larger mission of the University.

Certain types of field research might, however, require striking a particularly delicate balance among sometimes contrasting needs for free and informed consent, accuracy and objectivity, and distance and criticality. For instance, an anthropologist might wish to conduct interviews or participant observation with a neo-fascist group, and anticipate writing critically about this group’s views relative to, say, a liberal humanist framework. Such a scenario should not be understood as an exception to the general principle of respect for free and informed consent; nor, however, should it be understood as giving such a group veto power over a researcher’s right to be critical. Such issues should be carefully articulated in a protocol, and carefully navigated and negotiated in the field. Researchers should also be sensitive to the fact that some groups—such as aboriginal groups—might have a legitimate claim to want to work closely with researchers from

conception to completion, to help avoid “group harms” that researchers have contributed to historically.

The scenarios just described—i.e., in which a researcher directly interacts with participants through interviews and participant observation—should not be confused with naturalistic observation, in which the researcher typically does not directly interact with participants, and the types of data collected do not render individuals reasonably identifiable. Naturalistic observation is explicitly dealt with in the TCPS, Article 2.3 and related commentary, and the usual requirements for free and informed consent may legitimately be waived in cases that fall under those guidelines.

One other research method that is relevant in this connection is the so-called audit methodology, which presents a kind of limiting case scenario for exceptions to general guidelines for free and informed consent. For example, a sociologist or economist might use the audit methodology to directly respond to real employment or rental advertisements using fabricated applications that are designed to test for systemic biases such as racism or sexism, with no provision for free and informed consent. Such methods seem appropriate only in highly circumscribed scenarios, namely those in which privacy is not a legitimate expectation because public and commercial acts such as employment or rental advertisements are involved, and there are public regulations proscribing discriminatory activity (for a broader discussion of this issue, see, e.g., Riach & Rich, 2004).

### Deception and Debriefing

The remainder of this document will focus on research topics that entail less than full disclosure at the outset, and indeed that might require actively supplying participants with misleading information, but that can be managed by full disclosure in the form of a debriefing and re-consent option at the end. These are usually, although not necessarily exclusively, associated with experimental or laboratory research.

Some types of research might entail less than full disclosure at the outset, but not involve actively deceiving participants with misleading information. For instance, a researcher might truthfully explain to a participant the general topic of the research, but not be able to explain the specific focus of the study at the outset, at risk of altering the phenomenon in question. For example, a linguist might truthfully be able to explain at the outset that they are interested in discourse styles—but not be able to explain until later that their specific interest was in quotative uses of go and like (as in, He goes, “Uh oh”, and I’m like, “Whatever”) and its relation to speaker age. Such instances of less than full disclosure at the outset should, in general, be easily handled by providing participants with a full explanation at the end of the study, in the form of a verbal or written debriefing.

Research that involves active deception—that is, intentionally providing participants with misleading information over the course of the study—not surprisingly gives rise to more varied and complex ethics issues. The following examples are intended to give some sense of the range of potential scenarios and issues:

Example 1: A psychologist interested in the effects of various personality and social factors on eating behaviour might usher a participant into a “waiting room”, and provide her with a brief description including the sex, height, weight, and dieting habits of

“another participant” she will ostensibly be meeting for a study on “ice breaking”; in the meantime, the participant is asked to fill out a series of personality questionnaires, and provided with a tray of snacks. In fact, however, there is no “other” participant; rather, the actual focus of the study is the effect of participant personality and manipulated characteristics of the hypothetical “other” participant on the amount of food participants eat.

Example 2: A management researcher interested in the influence of religion, science, and politics on consumer decisions might present participants with quotes attributed—sometimes falsely—to real, well-known figures from these different fields, before testing whether the different quote attributions influence subsequent consumer decision making.

Example 3: An education researcher interested in the effects of stereotypes on academic performance might present participants with different “scientific” statistics regarding correlations between sex or race and performance in different academic subjects, before testing whether the priming of these stereotypes affects subsequent performance on different sub-sections of standard achievement tests.

Such examples clearly do not involve free and informed consent at the outset; they can, however, be managed with a debriefing and re-consent option at the end. The issues that a debriefing would need to address are somewhat complicated, however, and bear bringing out. For instance:

Privacy: A participant might reasonably feel that their privacy has been violated, if the researcher has collected information—e.g., how much food the participant ate—that the participant considers private, and not something they initially consented to. They might also feel that their privacy has been violated if the information collected is to be used to some other end than they would agree with—for instance, to pit science against religion.

Misinformation: If a deceptive scenario deliberately misrepresents facts about real individuals, groups, or states of affairs in the world, the researcher has a responsibility to set the record straight. This involves not only clearly identifying which elements of the study were outright fabrications or misrepresentations, but also providing a positive statement of actual facts. This could be particularly important if the topic relates to issues that participants might reasonably be expected to consider important.

Temporary and residual harm: If a deceptive scenario involves provision of false information that might be understood as applying to participants themselves, or to a group they belong to, the researcher should address and attempt to minimize any temporary negative arousal that might occur during the deception, such as frustration due to false failure feedback; the researcher should also address and attempt to minimize any residual misconceptions that might actually last beyond the period of the study itself. At least one study has shown, for example, that participants who received false failure feedback that was debunked during debriefing subsequently gave lower estimates of their actual and future performance on the task in comparison to participants who had received success feedback that was similarly debunked (Ross, Lepper & Hubbard, 1975).

To manage such issues, researchers should abide by the following guidelines, as suggested by the TCPS:

- Studies involving active deception should have a clear and defensible scholarly motivation—and each of the deceptive elements of the study should be *necessary*, with a clear rationale to back this up.
- Participants should be provided with a complete, plain-language *debriefing* that explains very frankly:
  - which elements of the study were deceptive
  - why these were necessary
  - how the study relates to a broader, important area of knowledge
  - once participants have been fully debriefed, they should be given the opportunity to re-consent to the use of their data, so it is clear that they have in fact given informed consent, and understand that they are otherwise free to withdraw their data
- Research involving active deception should be relatively low risk—which is to say, should not involve therapeutic interventions, or highly vulnerable groups and sensitive topics.

To give a sense of scenarios in which deception would be *inappropriate*, it is difficult to imagine a rationale that would make it seem reasonable to use deception with a highly vulnerable group such as children who have been abused. It is similarly difficult to imagine justifying deception in a study involving a highly sensitive topic such as individuals' experience of child abuse. It can also be difficult to justify deceptive study designs involving some combination of moderate group vulnerability and moderate topic sensitivity, as in a study of young children's frustration in response to false failure feedback. Such studies should be considered in terms of the University of Toronto's *risk matrix*, which treats group vulnerability and research risk each as low, medium or high. Combinations based on the matrix ratings that result in studies classified as "low-low", "low-medium", or "medium-low" can be expedited; however, deceptive study designs that threaten to go beyond these should be treated with considerable discretion by researchers and REBs. In particular, researchers and REBs should think of such designs in terms of underlying principles of respect for human dignity and respect for vulnerable persons, so that vulnerable groups and sensitive topics are not used as a means to some relatively inconsequential end.

It should be clear from the foregoing that if active deception is properly thought through and properly handled, it is not inherently unethical, as some individuals might presume; indeed, if good practices are followed, participants who have been properly debriefed should be unlikely to find fault with the study, or to wish to withdraw from it. If participants do have concerns, however, they should be in a position to contact the Research Ethics Board (REB), and the researcher should work with the REB to address any issues.

The following additional remarks are intended to draw attention to some common pitfalls, relative to good practices:

Initial consent: It is a good practice to make the initial consent process as informative as possible, keeping as much deception out of it as possible. To this end, it may make sense to mention explicitly that while it is not possible to fully explain the study in advance, participants will be provided with a full debriefing at the end.

*"neither arbitrary nor capricious"* (phrase used in the TCPS): Study designs involving active deception should be approached with the same intellectual rigour as any other

aspect of scholarship; they should not be taken as license for all manner of unnecessary fabrication. Indeed, sticking as closely to the truth as possible should be in everyone's interest because it makes the deceptive elements as straightforward as possible to introduce and remove. To think of this in terms of a slogan: do the minimum to deceive; do the maximum to debrief.

**Tone:** A sensitive and informative debriefing can go a long way to mitigating possible adverse effects of deception. Participants should not leave feeling they have been duped. Given the potential for loss of trust in the research community, and for residual misunderstanding and even harm, researchers should work diligently to ensure that debriefings really are received and absorbed. A common pitfall is for debriefings to be perfunctory and not particularly informative. For example, technical psychological jargon such as, "the experimental manipulation was . . ." or "the independent variable was . . ." can effectively bury the fact that some people were told something true and some people were told something false. Such indirect or jargony debriefings can, moreover, give the impression the researcher has something to hide, or is debriefing "through gritted teeth". Researchers should therefore make a special effort to use plain language with participants, to "come clean" about deceptive elements, and to situate the study in the larger area of interest.

**Relevance:** Researchers should be sensitive to the fact that the specific focus of a particular study might relate to issues that are actually of some personal concern to participants—such as stereotyping, psychological well-being, body image, or eating disorders. In situations where this is reasonably foreseeable—for example, given known base rates for such concerns among the population of interest—then it would be a good practice for the researcher to have on hand contact information for relevant services or general background information so that participants can pursue any larger questions or issues raised by the study.

**Training:** Given the complexity and sensitivity of issues surrounding deception and debriefing in research, supervisors should work closely with students to provide training and raise awareness regarding good practices for the use of these methods.