The Costs of Deception: Evidence From Psychology

by

Andreas Ortmann[a,b] *and* **Ralph Hertwig**[a,*]

[a]Center for Adaptive Behavior and Cognition
Max Planck Institute for Human Development, Berlin, GERMANY

[b]Center for Economic Research and Graduate Education
Charles University, Prague, CZECH REPUBLIC

*Correspondence: Andreas Ortmann, Ph.D.
[until Augst 2000]
Center for Adaptive Behavior and Cognition
Max-Planck-Institut fuer Bildungsforschung
Lentzeallee 94, D-14195 Berlin, GERMANY
ortmann@mpib-berlin.mpg.de
Tel: (011)(49)(30) 824 06 438 (office)
Center for Economic Research and Graduate Education
Charles University
Politickych veznu 7, 111 21 Prague 1, CZECH REPUBLIC
andreas.ortmann@cerge.cuni.cz
Tel: (011)(420)(2) 240 05 117 (office)
Alternatively:
(011)(49)(173) 65 05 191 (cellular, works for both Berlin and Prague)

Authors 'note: We would like to thank Valerie M. Chase, Gerd Gigerenzer, Anita Todd, Peter M. Todd, Tom Wallsten, David Weiss and participants at the ESA/Public Choice meetings 2000 in Charleston, S.C., USA, for many constructive comments.

Abstract

Recently, it has been argued that there is no evidence in social science research that deceiving subjects in an experiment leads to a significant loss of experimental control. Based on this assessment, experimental economists were counseled to lift their de facto prohibition against deception to capture its allegedly significant benefits. Previous reviews of the evidence, however, have not been systematic. Here we provide a systematic assessment of the methodological costs of deception, mostly drawing on evidence from psychology where deception, in areas such as social psychology, is wide-spread and frequent and has a long tradition.

We report two major results. First, the evidence suggests that deception is bound to generate suspicion and second-guessing of the research design and is likely to affect judgment and decision making of a non-negligible number of participants, thus at the minimum inducing unnecessary noise in the data and reducing significantly procedural regularity and replicability. Second, we find little evidence for reputational spillover effects hypothesized by a number of authors (e.g., Kelman, 1967; Davis and Holt, 1993).

We conjecture that the latter finding may be the result of institutional arrangements that psychologists have evolved in response to the very problem. There is, therefore, no guarantee that reputational spillover effects would not arise if deception were to become an accepted methodological tool in economics. Furthermore, allowing deception would be likely to slow down methodological innovation. We conclude that the prohibition of deception is a sensible convention that economists should not abandon.

Keywords: Experimental economics, deception, reputational spillover effects

JEL classification: C72; C91

1. Introduction

Among experimental economists, deception is generally taboo; exceptions can probably be counted on two hands.[1] Davis and Holt (1993, pp. 23-24) gave the following typical rationale for why researchers should avoid deceiving subjects:

> Most economists are very concerned about developing and maintaining a reputation among the student population for honesty in order to ensure that subject actions are motivated by the induced monetary rewards rather than by psychological reactions to suspected manipulation.

Those advocating the use of deception in experimental economics argue that there is little evidence from social science research that deception leads to such a loss of experimental control. Hence, they argue that economists 'long-standing and effective proscription of deception is unduly restrictive and prevents the capture of the potentially significant benefits of using it (e.g., and prominently, Bonetti, 1998, 1998a). Economists 'response to this argument has been swift but non-empirical (e.g., Hey, 1998; McDaniel and Starmer, 1998). Whether indeed the use of deception incurs methodological costs, however, is an issue that can be evaluated empirically.

[1] Bonetti (1998) and Hey (1998) enumerate the half dozen studies in economics that come to mind quickly. In a recent survey, we found that a representative sample of experimental economists estimated that, on average, they use deception in 0.17 of 10 experiments (MD = 0, SD .44, for details see Hertwig and Ortmann, *forthcoming*).

Although the de facto prohibition of deception in experimental economics makes such an evaluation an impossible enterprise, we can draw on data from psychology where the use of deception, in some areas, is wide-spread and frequently has a long tradition (Hertwig and Ortmann, 2000). While the research agendas, practices, and subject pools of psychologists and economists diverge significantly (Hertwig and Ortmann, *forthcoming*), experimenters in both disciplines do share research areas of common interest such as judgment and decision making and social psychology. Psychologists in these areas, in particular in social psychology, happen to use deception experiments extensively and intensively, inviting a closer look "across the border".

If methodological costs to the use of deception are found, should they be incurred? In psychology, answers to this question have been given from a consequentialist or a deontological point of view. Those favoring the deontological point of view argue that the costs of any violation of subjects' human rights are prohibitive (e.g., Baumrind, 1964, 1979, 1985). This view effectively outlaws the use of deception whatever its benefits. Those favoring the consequentialist approach have argued that the costs of violating subjects' human rights (not considered prohibitive) have to be weighted against the costs of not being able to test theories and/or to explore experimentally important social policy issues. This view motivates the kind of cost-benefit calculus that the American Psychological Association has adopted in theory (albeit not in practice[2]) and that we will pursue here. This does not mean that we intend to take sides in

[2] The APA guidelines admonish researchers to employ deception as a last-resort strategy only. However, the high rate of deception experiments in areas such as social psychology strongly suggests that deception is not treated as a last-resort strategy. Deception, in many areas of psychology, is still considered a cavalier delict that "unrepentant deceivers" (as one psychologist called himself in a communication to one of us) find easy to rationalize with the elegance of a research design. In Hertwig and Ortmann (2000) -- written for a psychology audience and a

the debate between consequentialists and deontologists. It also does not mean that we consider the deontological arguments to be irrelevant. Our view is, simply, that agreement on the merits of deception is, if at all, more likely to be reached over issues such as the costs of reputational spillover effects. The purpose of this paper is therefore to empirically assess the methodological costs of deception.

The paper is organized as follows: Section Two provides a definition of deception and sketches the extent of its use in areas such as social psychology. Section Three discusses the game-theoretic structure of the social situation "experiment." Building on the game-theoretic framework, we review in Section Four evidence on "direct" and "indirect" effects of deception experiments on both those who use it and others. In Sections Five and Six we summarize our findings, discuss contrary views, and examine whether the benefits that economists might capture could justify the costs of using deception.

2. How wide-spread, and frequent, is the use of deception in psychology?

Before we answer this question, a clarification of what constitutes deception is in order. Psychologists and economists agree that intentional misrepresentation of the purpose of an experiment, or any aspect of it, constitutes deception. For instance, psychologists Adair, Dushenko and Lindsay (1985) stated: "Deception was defined as the provision of information that actively misled subjects regarding some aspect of the study" (p. 62). And Nicks, Korn, and

---

companion paper of sorts -- we have advocated an incentive-compatible mechanism that we believe might have a chance to reduce the amount of deception experiments in psychology.

Mainieri (1997) defined deception as "an explicit misstatement of fact"(p. 70). Along these lines, Hey (1998, p. 397) stated succinctly, "There is a world of difference between not telling subjects things and telling them the wrong things. *The latter is deception, the former is not.*"

Or is it? What if not telling participants things leads them to invoke certain default assumptions? One default assumption participants may reasonably have is that the initial interaction with the experimenter (upon entering the laboratory) is not yet part of the experiment. Another plausible default assumption participants may have is that the participants in the experiment are the other subjects in the room (see Gross and Fleming, 1982; Sieber, Iannuzzo, and Rodriguez, 1995, p. 72).

Clearly, we are getting into tricky territory quickly. While most researchers in both disciplines seem to agree that it is not feasible, and desirable, to acquaint participants in advance with all aspects of the research being conducted (e.g., its hypotheses and the full range of experimental conditions)[3], absence of full disclosure and "economy with the truth"can violate default assumptions and therefore mislead participants. We do not know what the effect of violations of default assumptions is. However, we propose that violations of default assumptions, whether resultant from a sin of commission or omission and whether intentional or not, can generate suspicion, induce second-guessing of the research design, and have effects on

[3] Even Baumrind, whose persistent critique of deceptive practices (e.g., Baumrind, 1964, 1979, 1985) made her the doyenne of "The Prohibitionists"(Bonetti 1998, p. 379) in psychology, suggested that "absence of full disclosure does not constitute intentional deception"(Baumrind, 1985, p. 165).

participants ' judgment and decision making similar to those produced by deception. Hence violations of default assumptions should be avoided.

How wide-spread then, and frequent, is the use of deception in psychology? Take the highest-ranked journal in social psychology, the Journal of Personality and Social Psychology (JPSP), and its predecessor, the Journal of Abnormal and Social Psychology as an illustration.[4] After a sharp upswing during the sixties when the percentage of deception studies tripled from 16% in 1961 to 47% in 1971, the use of deception continued to increase throughout the seventies, reaching its high in 1979 (59%) before it dropped to 50% in 1983 (Adair et al. 1985). Since then it has fluctuated between 31% and 47% (1986: 32%; 1992: 47%; 1994: 31%; 1996: 42%; as reported in Adair et al., 1985; Sieber et al.,1995; Nicks et al., 1997; and Epley and Huff, 1998) While some of these fluctuations may reflect different definitions of what constitutes deception (e.g., compare the more inclusive criteria employed by Sieber et al. with the criteria used by Nicks et al.), a conservative estimate would be that every third study published in JPSP in the 1990s employed deception. In other social psychological journals, for instance, Journal of Experimental Social Psychology, the proportion is even higher (Adair et al., 1985; Nicks et al., 1997). The widespread and frequent use of deception in social psychology in recent decades contrasts markedly with its decidedly more selective use in the 1950s and earlier (Adair et al., 1985). Despite the guidelines of the APA, deception seems to have become a first-resort strategy

[4] We found 14 studies that analyzed the use of deception across a wide range of journals and areas of study (e.g., Toy, Olsen, and Wright, 1989). We use JPSP as illustration because the data for this journal are the most comprehensive and recent. Arguably, this journal also covers an area where there is significant overlap in research interests of economists and psychologists.

rather than the last-resort strategy that it is meant to be.[5] As we have argued elsewhere (Ortmann and Hertwig, 1997, 1998; Hertwig and Ortmann, 2000), the reasons for this may well have to do with the fact that psychologists –being allowed to weigh their own private benefits of using deception against the public costs –are bound to fall prey to the implicit moral hazard problem.

3. The structure of the social situation "experiment"

To structure the presentation of the empirical (experimental) evidence, it is useful to first clarify the underlying strategic situation in which both experimenters and participants find themselves. We suggest that the social situation "experiment" has the complicated structure of a multilevel game (e.g., Hausken, 1995, 1995a). As such it is, for one, a complex game which is played inter-group (between experimentalists and participants) and intra-group (between participants or between experimentalists). Furthermore, as Schultz (1969) observed, the relationship between participants and experimentalists "has some of the characteristics of a

[5] Although deception is still widely used, ethical guidelines for research have become stricter (for a short history of the "ten commandments of the APA" see Rosnow and Rosenthal, 1997, chapter 6). As a consequence, the profession seems to have experienced few spectacular instances of deception. Rosnow and Rosenthal (1997), for instance, concluded that, "many of the seminal studies that were conducted then would be impossible today (e.g., Milgram's obedience studies)" (p. 114). We agree with this conclusion, notwithstanding evidence that emotional distress caused by less severe deception practices can still be substantial (e.g., Oliansky, 1991).

superior-subordinate one . . . Perhaps the only other such one-sided relationships are those of parent and child, physician and patient, or drill sergeant and trainee."(p. 221) Argyris (1968) similarly argued that "[subjects] are now beginning to behave like lower level employees in companies"(p. 187). In other words, the already complex multilevel game "experiment"is complicated through strong hierarchical components. These may invite psychological reactions per se that may have little to do with the experimental scenario or the reputation of the lab among the student population. It is likely that these problems are more prominent if participation is not voluntary.

3.1. *The inter-group game between participants and experimental economists.* The choices available to experimenters are those of an agent squaring off with a principal in a one-off one-sided social dilemma or reputational game (Ortmann and Colander, 1997; Kreps, 1990). Both agent and principal can either contribute their respective assets (honesty for the agent, trust for the principal) or withhold them. Specifically, the agent (experimenter) can choose either to deceive participants (the principals) or to be truthful about the setting and purpose of the experiment. The principal (participant), in turn, can choose either to trust the experimenter or to doubt the experimenter's claims. The game-theoretic predictions for this game are, dependent on the parameterization, clear-cut. The agent will defect (at least with some probability). The principal, anticipating the defection, will doubt the experimenter's claims about the experiments alleged purpose and nature of the experiment (at least with some probability).[6]

[6] See Ortmann and Colander (1997) for two parameterizations: for one defection is the (weakly) dominant strategy, for the other defection is in the support of the mixed-strategy equilibrium.

The interaction between agent and principal, of course, may not be a one-off game. Participants (principals) may partake repeatedly in experiments, quite possibly by the same experimenter. If they are experienced in that sense, then that experience may bear "directly" on their expectation of the experimenter's action choice. Even if inexperienced in that sense, however, participants may come with expectations of being deceived. Davis and Holt (1993, p. 24) sketch a possible transmission mechanism:

> Subjects may suspect deception if it is present. Moreover, even if subjects fail to detect deception within a session, it may jeopardize future experiments if the subjects ever find out that they were deceived and report this information to their friends.

Such "indirect" effects can come about through campus scuttlebutt, but also news reports, (e.g., the International Herald Tribune (September 18, 1996) on the Nisbett and Cohen study of "The Bellicose Southern Male[7]"), articles documenting deception and participants' reactions to it (e.g., Oliansky, 1991), and – possibly most important – classic undergraduate teaching and textbook examples (Milgram, 1963, 1964; Asch, 1955, 1956). Importantly, indirect effects transform seemingly one-shot principal-agent games into repeated games which, of course, have very different equilibrium predictions in the kind of social dilemma or reputational game we investigate here.

[7] A couple of years after the study came out, one of us, while taking in the spectacular vista of the Mohegan Island coast (Maine), overheard one hiker tell another one about that very experiment.

3.2. *The intra-group game between experimental participants.* In experiments, participants may interact with other participants. In games such as the trust game (Berg, Dickhaut, and McCabe, 1995; Ortmann, Fitzgerald, and Boeing, *forthcoming*), they may feel "deceived "by decisions of other participants that do not correspond to their expectations. While this interaction clearly does not constitute deception on the part of experimenters, they might want to anticipate that participants may blame them for other participants 'choices or may doubt that decisions were made by other participants.[8] In our view, the only remedy against such inferences is to make the experimental design and implementation as transparent as possible (including potential consequences) and to maintain a reputation for not tampering with any aspect of the experiment.

3.3. *The intra-group game between experimental economists.* Participants 'expectations that they will not be deceived (i.e., honesty on the part of the experimenter) is a common good. If the experimenter chooses to deceive participants, and is found out or suspected of employing deceptive practices, then the common good might get depleted. Whether indeed it will be

[8] In the replication of the results by Berg and her colleagues (1985), Ortmann et al. *(forthcoming)* repeatedly encountered participants who were upset about the fact that other participants had not reciprocated their investments and, in one case, angrily questioned one of the experimenters afterwards about the ethics of what they were doing (i.e., giving participants "the opportunity to exploit others "). Such reactions have to be taken seriously in settings such as small liberal arts college where they become, more or less inevitably, fodder for the gossip mill.

depleted is a function of the reaction of the participants. As we will see presently, a non-negligible fraction of participants will react in a hostile, or at least non-cooperative, manner. Their reaction, however, may or may not generalize. If it generalizes, their uncooperative behavior will be directed against all experimenters in general. If it does not generalize, uncooperative behavior will be limited to the experimenter, or her or his lab. In the second case the costs of deception are internalized. In the first case, however, the costs of deception create externalities. Importantly, as already mentioned, externalities can be generated through both direct experience and indirect experience with deception.

Clearly, direct and indirect effects are conditioned on a number of contingencies, including the question to what extent participants will distinguish individual reputations from collective reputations for deception experiments. The importance of collective reputation has been suggested by Tirole (1996) who defined it as the average quality produced by a group of firms which are similar to a reference firm. Landon and Smith (1998), in a titillating study, show that both individual and collective reputation of Bordeaux red wines matter. However, the marginal impact of expected quality (i.e., collective reputation) on price is approximately 20 times greater than that of current quality (i.e., individual reputation). Related are results from research in social stereotyping that suggest that individuating information such as specifics about a person often gets ignored due to potent base rate beliefs (e.g., Locksley, Hepburn, and Ortiz, 1982; Koehler, 1996). If these results were to hold for professional reputations, then an individual researcher's attempt to avoid deception and establish a good reputation for her or his lab would clearly be doomed.

The game-theoretic analysis of the interaction between participants and experimenters

suggests that any discussion of the possible negative effects of deception needs to distinguish four classes of consequences. The first two classes are: the effects of direct experiences (i.e. firsthand experiences with deception) on individual experimenters and on the profession as such. The second two classes are: the effects of indirect experiences (i.e. vicarious experiences with deception) on individual experiences and on the profession in general. Figure 1 shows these four contingencies. Note that only in the upper-left cell do we encounter the possibility of immediate cognitive-emotional and behavioral responses in a given experiment. In contrast, all other cells describe time-delayed responses in various forms.

-- Figure 1 about here --

In the following section we will discuss the available evidence from psychology, as well as relevant experimental evidence from economics, one contingency at a time. Within each class, we will distinguish whether deception affects emotion and cognition (i.e., things such as feelings, attitudes, beliefs, and expectations) and/or experimental performance. This distinction does not necessarily define mutually exclusive categories; indeed some people argue that emotions are behavioral responses, or at least immediately translate into such. As we will show, evidence that researchers have marshalled of the effects of deception experiments is rather mixed. We first describe the evidence in Section 4; possible reasons for the mixed picture will be discussed in section 5.

4. Consequences of the use of deception: Evidence

4.1. *How participants 'direct experiences with deception translates: Individualized effects*

Emotional and cognitive responses. A series of authors have concluded that during and/or

after the conclusion of deception experiments, participants display negative emotions. Among others, Fisher and Fyrberg (1994) reported that the majority of their students believed that participants in various published deception studies must have felt embarrassed, sad, or uncomfortable. Finney (1987) found that deceived participants in his conformity experiments, compared to non-deceived participants, that his participants believed less in the value of the research, reported less confidence in their judgment, and more tension during the experiment. Allen (1983), in prisoner's dilemma games, found that participants who had been deceived during the session "rated the experiment as worthless, were annoyed with the experiment, and would not recommend the experiment to a friend" (p. 899; see also Straits, Wuebben, and Majka, 1972); others did not. Oliansky (1991) observed that both participants and research assistants/confederates exhibited severe negative emotions[9].

[9] Oliansky related his experiences as a confederate/research assistant whose job it was to trick participants into believing that they could affect another person's emotional well-being. In actuality, the experiment was designed to test how participants' feelings were affected by their perception of their ability to help another person in emotional distress. Oliansky discussed his own doubts and guilt as a confederate/research assistant; he also discussed the angry responses of the significant number of participants who felt duped and the responses of others who questioned the alleged purpose of the experiment from the beginning.

In stark contrast to these reports of negative emotional and cognitive responses, Christensen (1988) summed up his review of research on the impact of deception on participants as follows: "This review ... has consistently revealed that research participants do not perceive that they are harmed and do not seem to mind being misled. In fact, evidence exists suggesting that deception experiments are more enjoyable and beneficial than non-deception experiments." (p. 668)[10] Interestingly, Finney (1987) also found that his deceived subjects experienced less boredom with the experiment than his non-deceived subjects, although the deceived subjects reported more tension, less confidence in their own judgments, and more doubts about the value of the research. Noting that records of participants' reactions to deception often draw on surveys or people imagining taking part in experiments and then reporting their imagined reactions, Aitkenhead and Dordoy (1985) designed an experiment that crossed active and passive (role-playing) participants and examined the effects of deception, physical discomfort (stress), and experimenter considerateness. They found that in contrast to the imagined reactions of role-playing participants, "[active] subjects have, on the whole, positive reactions to the research, and are generally tolerant towards being deceived, being caused physical discomfort (even pain), and being treated inconsiderately by the experimenter"(p. 303) . It is noteworthy that the experiments reported in Finney (1987) and Aitkenhead and Dordoy (1985) involved minor forms of deception, physical discomfort, and/or inconsiderateness[11], which may have contributed

[10] Curiously, his review did not contain a previous article (Christensen 1977) in which the author found that "subjects who perceive that their behavior is being manipulated will tend to resist this influence and exhibit behavior reflective of the so-called negative subject."(p. 399)

[11] Aitkenhead and Dordoy (1985) had participants solve five-letter anagrams. Deceived participants were told that the aim of the experiment was to discover how quickly anagrams

to the fact that participants were not bothered much by experimenters' use of deception.

The seemingly conflicting evidence prompts the following five observations and comments: First, the fact that many psychological experiments are of a "mundane, repetitive nature" (Sharpe, Adair, and Roese, 1992, p. 589; see also Coulter, 1986) cannot justify the use of deception. Rather, it should spur experimenters to make experiments more interesting and provide participants with incentives to perform well in them (e.g., Argyris, 1968; Hertwig and Ortmann, *forthcoming*; Hertwig and Ortmann, 2000). Second, to the extent that (the possibility of) deception makes a tedious, self-evident, or otherwise less interesting study less so, it may well trigger curiosity, more perceived educational benefit, and higher overall ratings of the research program. However, these affective responses (whether they are negative or positive) are experimental artifacts. As such they may undermine economists' saliency tenet (Smith, 1982), especially if the pay-off function is flat. Third, deception (as used in Aitkenhead and Dordoy, 1985) is not deception (as used in Finney 1987) is not deception (as used in Milgram, 1963, 1964; or, Oliansky, 1991). That is, whether deception lies within a participant's "comfort zone"

---

could be solved when participants were either relatively stressed or relatively unstressed. The high-stress condition required participants to solve three difficult anagrams while holding an ice cube; the low-stress condition required participants to solve three easy anagrams (no ice cube). Finney's (1987) experiment was a conformity experiment that exposed participants to incorrect line-length judgments of a group of confederates in the tradition of Asch (1955, 1956). The Aitkenhead and Dordoy participants were not paid; Finney's participants were offered extra credit points that were, however, mostly inconsequential.

(Gerdes, 1987), is a function of issues such as the nature and severity of deception, the methods of debriefing, the recruitment mode, etc.

Fourth, the findings regarding participants' apparently positive feelings and attitudes are not as unequivocal as they were sometimes presented. For an illustration, consider the following study prominently featured by Christensen (1988, p. 668). Smith and Richardson (1983) concluded that "those participants who had been deceived evaluated their experience more positively than those who had not participated in deception experiments" (p. 1075). As Rubin (1985) pointed out, the same authors also reported that 20% of the participants in their survey (undergraduate students in introductory psychology classes) reported they had been in experiments that "caused them to feel very nervous, humiliated, excessive physical discomfort, very wrongly deceived and/or very angry" (p. 1078). These observations are relevant because there was a correlation (albeit not perfect) between those feelings and participation in a deception experiment.

Fifth, and possibly most important for the present purpose, even if people had positive emotions, it would not preclude the possibility of them becoming suspicious, arguably an emotional-cognitive response. As Kelman (1967) put it, widespread deception is likely to lead to a situation in which participants "may not know the exact purpose of the particular experiment in which they are participating, but at least they know, typically, that it is *not* what the experimenter says it is" (p. 6). In the next section, we explore how prevalent this kind of suspicion is among participants in psychology experiments.

We conclude that even if one accepts Christensen's (1988) questionable contention that participants in deception experiments actually enjoy and believe to benefit from them, one would

still face the question of to what extent such emotional and cognitive responses would undermine procedural regularity and add noise to the experimental data-generating process.

The prevalence of suspicion. Do participants typically assume that the purpose of an experiment is not what the experimenter says it is? To get a more systematic idea of the prevalence of suspicion among experimental participants in psychology, we conducted a literature search in the PsycINFO/PsycLIT database, which covers the academic literature in psychology and related disciplines, including sociology and education, in the period between 1887 and July 1999 (when our searches were conducted). We searched for specific keywords (listed below) in titles and abstracts. We also included all studies cited in Bonetti (1998) who concluded from his sample that "deception does not appear to jeopardize future experiments' or 'contaminate the subject pool'" (p. 389). Finally, we looked up the crucial studies cited in the articles found and included them if they seemed relevant for the issue at hand.

Specifically, we searched for the key word deception in combination with suspicion (and its variants such as "suspicious, suspiciousness, suspicions"). This search uncovered two systematic reviews of the social psychology literature (Stricker, 1967; Stang, 1976) that examined the proportions of participants who are suspicious. In a sample of 88 deception studies from four leading social psychology journals, Stricker (1967) found that only 16 studies attempted to quantify the degree of participants' suspicion of deception. In this subset, the median percentage of "suspicious" participants was 4%, with a range from 0% to 23%. Stricker, Messick, and Jackson (1969, p. 345) later suggested that "It seems very likely that the overall rate of actual suspicion in these 16 studies was seriously underestimated" due to the inadequacy of the suspicion criteria used. Using different criteria, later studies such as Glinski, Glinski, and Slatin (1970), Ettinger,

Mariono, Endler, Geller, and Natziuk (1971), Endler and Hartley (1973), Geller and Endler (1973), Geller, Endler, and Wiesenthal (1973), Wiesenthal, Endler, and Geller (1973), Willis and Willis (1970), Rubin and Moore (1971), Adair (1972), and Stang (1976) found indeed much larger numbers of suspicious participants: typically between one third and two thirds of participants.

Through an exhaustive search of the conformity literature, Stang (1976) identified 21 studies that reported the percentage of participants who were classified as suspicious.[12] Interestingly, Stang found a dramatic increase of suspicion, especially in the second half of the 1960's that seems to track closely the dramatic increase of deception experiments during that period and the decreased credibility of social psychologists and/or increase in participant sophistication and decrease in participant trust.[13]

[12] Typically, the classification is done on the basis of post-experimental interviews that prompted participant answers to questions such as "Do you feel this experiment was deceptive (involved lying) in any way?"(Geller and Endler, 1973, p. 49).

[13] Stang (1976, p. 355) mentions other explanations for the increase in suspicion, namely decreased quality of the cover story and measurement artifacts. Both strike us as unlikely.

Stang (1976, p. 360) claimed that participants in conformity experiments who report suspicions commonly get excluded from the data analysis.[14] (In Stang's 1976 study, 20% of participants were excluded based on self-reported suspicions.) The problem is, as we will presently see, it is quite possible that this number systematically under-estimates the true degree of suspiciousness since participants often will not truthfully reveal their suspicions or knowledge (e.g., Argyris 1968; Newberry 1973; Oliansky 1991; Taylor and Shepperd 1996). Suspicion, in other words, can quickly turn good-subject behavior into bad-subject behavior.

In light of the high number of suspicious participants in conformity experiments, note that Epstein, Suedfeld, and Silverstein (1973) observe that deception is, next to a physical danger to the participant, the most frequently mentioned reason to withdraw from an experiment. Thus, cognitive-emotional responses may translate into a particular form of behavioral response -- physical withdrawal during the experiment -- that from the view of the scientific hypothesis under consideration is not meaningful.

Behavioral responses/performance. There are two strategies of analyzing the effects of suspicion on experimental performance. Experimenters can engender participants' suspicion from the outset and study their subsequent performance as a function of it. Or, the experimenter can record suspicion after participants perform the experimental task. In what follows we explore each of these strategies.

Experimentally induced suspicion. To find such studies, we did another search following the

[14] Conformity experiments, however, are by now means the only areas of research where suspicious participants (self-reported) get excluded (e.g., Sagarin, Rhoads, and Cialdini, 1998).

procedure outlined earlier. Specifically, we used the search term deception in combination with prebriefing, or forewarning. We found eight studies (for details see Table 4 in Hertwig & Ortmann 2000). Participants' knowledge and thus suspicion ranged from relatively neutral forewarning about experimental procedures in general (e.g., Allen, 1983, p. 901: "In a few experiments it is necessary for experimenters to deceive subjects concerning some elements of the experiment") to confederate tip-offs (e.g., Levy 1967) and disclosure that deception would occur during the experiment (e.g. Finney, 1987).

To quantify the reported findings, Hertwig and Ortmann (2000) calculated effect size measures where possible.[15] The results, once again, were mixed. Some of the eight studies that we identified found no effects and some found large effects. Nevertheless, a trend is discernible. When participants received detailed tip-offs about the true purpose of the experiment (e.g., Levy, 1967; Turner and Simons, 1974), were explicitly told that they would be deceived (Finney, 1987), or explicitly acknowledged awareness of experimental manipulation (Golding and Lichtenstein, 1970), suspicion altered experimental performance (albeit not necessarily on all dependent measures). In contrast, when participants were merely informed that some kind of

[15] The effect size measure we used is eta. It is defined as the square root of the proportion of variance accounted for (Rosnow and Rosenthal, 1991), and is identical to the Pearson product-moment correlation coefficient when df is 1, as is the case when two conditions are compared (which was the case in most cases where we calculated eta). According to Cohen's (1988) classification of effect sizes, a value of eta of .1, .3, and .5 constitutes a small, medium, and large effect size, respectively.

deception might happen (Allen, 1983; Finney, 1987; Wiener and Erker, 1986) or were told the purpose of the study (without indicating the possibility of deception, Gallo, Smith, and Mumford, 1973), then their performance did not differ from that of control participants not given this information (but see Spinner, Adair, and Barnes, 1977)

There are two interpretations of these results. For one, one could conclude that specific and certain expectations about deception alter experimental performance, whereas general and uncertain anticipation of deception does not. Finney (1987), however, had another suggestion for why uncertain anticipation ("may be deceived") did not seem to differ from the control condition (with no information). In his view, "one might speculate that this [general and uncertain] information merely reaffirms subjects' prior belief that deception may occur in an experiment and, therefore, causes no change in their anticipation" (p. 45). If indeed this general uncertain information only reaffirms the prior belief, it is not surprising that the experimental and the control groups (which shares the same priors) do not differ.

Suspicion recorded after participants performed the experimental task. An alternative, admittedly imperfect, way to assess participants' suspicion is to ask them after the experimental task (but before the debriefing) whether they had any suspicions. Our search turned up 14 studies that studied behavior as a function of suspicion. All of these studies were concerned with conformity behavior. In 10 of the 14 studies, suspicious participants conformed less than unsuspicious participants. In the 4 other studies (Chipman, 1966; Willis and Willis, 1970; Endler, Wiesenthal, and Geller, 1972; Wiesenthal et al., 1973), the suspicion did not significantly change the amount of conformity behavior. No study reported that suspicion produced greater conformity. For 9 of the 10 studies (those in which the necessary information was given) we

could calculate eta: the reduction in conformity due to suspicion was of medium to large effect size (for details see Table 3 in Hertwig and Ortmann 2000). To conclude, the systematic sample that we investigated suggests strongly that certain deception experiments significantly affect participants' behavior. To the extent that conformity experiments explore people's compliance with social pressure, and adherence to social norms, and to the extent that social norms have been identified as an important control variable in a variety of experimental settings (e.g., Hoffman, McCabe, and Smith, 1996; Ortmann and Tichy, 1999), using deception clearly can undermine procedural regularity.

Kinds of cooperativeness. Do noncooperative agents (experimenters) provoke noncooperativeness on the part of principals (participants)? Another form of withdrawal stemming from cognitive-emotional responses to (suspected) deception is participants' attempts to beat experimenters at their own game. Summarizing evidence that participants were often openly hostile toward experimenters and mandatory participation requirements, and with explicit reference to participants' responses to deceptive practices, Schultz (1969) suggested "that such attitudes (....) can influence the subjects' performance in a number of kinds of experimental tasks seems eminently possible"(p. 220). Schultz concluded that "perhaps it is the researchers who are being deceived"(p. 220). Similarly, Argyris (1968) presented the results of a formal evaluation of the basic psychology course in one major university by nearly 600 undergraduates:

> They were given three topics from which they could choose one to evaluate thoroughly. ... an overwhelming majority of the students focused on the requirement, in the course, that they had to participate as subjects. The students were very critical, mistrustful, and hostile to the requirement. In many cases, they identified how they expressed their pent-up feelings by "beating the researcher" in such a way that he never found out (an activity frequently observed among frustrated employees). (p. 188; see also Ninner and

Handelsman, 1992).[16]

Indeed, other researchers found that participants were reluctant to admit that they had prior information about the experiment. Newberry (1973), for instance, reported two experiments in which participants received information about the experiment from a confederate and were later asked by the experimenters if they had prior information. The proportion of the participants who lied about the fact that they possessed prior information varied from approximately, .8 to .3 in various conditions. Taylor and Shepperd (1996) described an experiment in which they used deception to study the effectiveness of conventional debriefing procedures for detecting suspicion of deception among research participants. Notwithstanding the explicit instructions of the experimenter not to communicate while he left the room on a pretext, Taylor and Shepperd found that participants did communicate with each other. By doing so, participants found out that deception was involved in the experiment -- a discovery that they did not reveal during debriefing. Taylor and Shepperd (1996) concluded that "our observation suggests that participants may fail to supply, and may even withhold, information that is crucial to evaluating whether the procedures proved a valid test of the hypothesis "(p.

[16] In marked contrast to most psychologists, economists recruit their students in more or less randomly determined classes, through flyers, or E-mail; they also typically use performance-dependent payments. Thus, it is probably safe to assume that participation is voluntary. For more on the differences between experimental economists and experimental psychologists, see Hertwig and Ortmann (*forthcoming)*; see also Ball and Cech (1996).

887).

In brief, participants tend not to admit that they had privileged knowledge (e.g., Newberry 1973, but see also Levy 1967; Altemeyer, 1971; Taylor and Shepperd, 1996) and sometimes disclose information to other participants although they earlier agreed to withhold it (e.g., Wuebben, 1967; Lichtenstein, 1970; Lipton and Garza, 1978; but see Aronson, 1966). Both observations surely undermine the assumption that participants are cooperative (e.g., Orne, 1962; Fillenbaum, 1966; Broeder, 1998) and the hope that reputational spill-over effects can be limited. The result, however, does not necessarily imply that experimenters' duplicity causes participants to sabotage psychological research (as proposed by Schultz, 1969). To draw this conclusion we would need to know the baseline of cooperative behavior among participants. Unfortunately, we know relatively little about the base rate of cooperative behavior in psychological experiments. Results from experimental economics may be used as proxy. The well-known experimental results on public good provision and two-sided and one-sided prisoner's dilemma games (e.g., Ledyard, 1995; Ortmann and Colander, 1997) suggest that a substantial number of participants react adversely to non-cooperative behavior. For example, Ortmann and Tichy (1999) found that in repeated prisoner's dilemma games with turnpike matching protocol, 50% of the participants who get "burned" switch to non-cooperative behavior. Since, in every round, participants are matched with a new participant whose prior decision history they do not know, this defection rate is very likely a lower bound. Recent evidence from a once-repeated trust game confirms the conjecture. Dickhaut, Hubbard, and McCabe (1995) observed that participants often give others the benefit of the doubt for the first round. If, however, their trust gets violated, they typically make sure it does not happen a second time. As the title of their

manuscript states succinctly, "fool me once, shame on you, fool me twice, shame on me." Interestingly, this result is very much in line with results in Christensen (1977) who found that participants with prior manipulative experimental experience resisted verbal conditioning.

4.2. *How participants' direct experiences with deception translates: Generalized effects*

> "It is believed by many undergraduates that psychologists are intentionally deceptive in most experiments." (Ledyard 1995, p. 134)

Emotional and cognitive responses. Cook, Bean, Calder, Frey, Krovetz, and Reisman (1970, p. 189) found that participants with a history of deception studies considered experiments to be less scientific and less valuable and reported caring less about understanding and following experimental instructions. They also found that such participants were more suspicious of the truthfulness of experimenters. Rubin and Moore (1971) found a positive relation between number of psychology courses taken and suspicion (but no relation between number of experiments and suspicion). Higbee (1978) found a strong negative relation between training in psychology and perception of psychologists' truthfulness.[17]

To assess the impact of the use of deception on participants' generalized expectations systematically, we conducted yet another literature search. To our surprise, only few studies seem to have studied participants' expectations systematically. This time we searched for the key word "deception" in combination with "expectation(s)" in all titles and abstracts and found a grand total of five studies that explicitly investigated participants' expectations regarding psychology

[17] This is a nice study with unusually clean results discussed in more detail presently. Unfortunately, Higbee didn't control for the number of deception experiments that those interviewed participated in.

experiments, such as the anticipated truthfulness of information provided. Among these studies, the effect sizes ranged from small to medium-large. (For details see Table 2 in Hertwig and Ortmann, 2000.) It is interesting to note that the most recent, and in our view most careful, studies - those by Epley and Huff (1998) and by Krupat and Garonzik (1994) - have larger effect sizes than the earlier investigations. In what follows, we briefly describe these two studies.

Epley and Huff (1998) experimentally studied suspicion of participants as a result of deception. They gave participants a reading comprehension task plus fabricated performance feedback. At the end of this first session, participants were given either a full debriefing (explaining the deceptive nature of the experiment) or partial debriefing (not explaining the deception part). Response to the experiment was measured through a questionnaire immediately after the debriefing. Deception had an effect on one measure of suspicion ("As a participant in future research, I will not be suspicious of any information presented by the experimenter.") but not the other ("Psychologists are trustworthy and honest."). In other words, Epley and Huff found that suspicion of experiments and suspicion of the whole profession are not significantly correlated. Not surprisingly, employing deception will increase suspicion concerning future experiments among those who are made aware of the deception.[18] (This latter result is important because the APA demands that experimenters who employ deception must debrief participants

[18] These results are, maybe, not surprising. The participants, after all, were psychology students: disagreement with the statement could easily have created cognitive dissonance: perceiving psychologists to be dishonest and not trustworthy means to perceive one's own future profession to be a questionable enterprise. Moreover, experimental psychologists represent only a fraction of of the general category "psychologists." Last but not least, those who had been fully debriefed could rationalize their belief since the experimenters had proven themselves trustworthy and honest by revealing the deceptive nature of the experiment.

afterwards. Clearly, this well-meaning imperative has -- as suggested by our game-theoretic analysis of the social situation "experiment"-- the potential to backfire.[19])

In their Study 1, Krupat and Garonzik (1994) asked 255 psychology students from introductory-, intermediate-. and advanced-level courses to answer an 11-item questionnaire. The introduction to all items was identical: "If you were asked to participate in a psychology study, would you expect that . . . ." The key item continued "you will be misled or deceived in some way during the course of the study." Controlling for psychology courses taken and experimental experience as participants, Krupat and Garonzik observed that even one deception experience is sufficient to make students expect that they will be misled or deceived in other studies. These findings are roughly in line with those of Epley and Huff (1998). In contrast, psychology courses taken and experimental experience as participants had no significant impact.

The results by Krupat and Garonzik (1994) and Epley and Huff (1998) seem to contradict earlier results by Sharpe et al. (1992). These authors included one item in their questionnaire that referred to deception, "Experiments in psychology almost always involve deception." They observed that their three participant samples disagreed with this statement. However, two of their participant samples had never participated in a psychological study prior to the survey, and only 29% percent of the participants in the third had encountered at least one deception experiment. In fact, only 12% of the experiments in which those 29% of the participants participated involved deception, and "relatively mild forms of deception"(p. 588) to boot. In

[19] Toy, Olsen, and Wright (1989) report that debriefing happens rarely in research in marketing and conjecture that experimenters thus try to avoid the very problem.

addition, psychology experiments do not "almost always" involve deception. Thus, participants may have the veridicial expectation that deception is used but not necessarily "almost always". In addition, Sharpe et al. (1992) found that after participating in a number of experiments, participants "reported a somewhat negative attitude toward psychological research" (p. 585). Krupat and Garonzik explained the discrepancy between their and Sharpe et al.'s results as a consequence of Sharpe et al. being "less clear and complete on the specific role of prior deception" (p. 219) and having worded their items in a general way instead of in "a person-specific and action-oriented manner" ("you will be misled . . . ") (p. 219).

In sum, the results suggest that firsthand experience with deception increases participants' expectations of being deceived (psychologists' truthfulness) without much affecting their beliefs about psychologists' trustworthiness in general.

Behavioral responses/performance. To study systematically the issue of how exactly experience with deception experiments translates into behavioral responses in other future experiments, we used, in yet another literature search in the PsycINFO/PsycLIT database, the term deception in combination with experimental history. We found a total of nine studies (for details see Table 5 in Hertwig and Ortmann 2000). The results suggested that firsthand experience with deception or manipulation affects performance in future experiments, while mere disclosure of the possibility of deception in psychological experiments does not (Cook and Perrin, 1971; Christensen, 1977, Experiments 1 and 2). Second, Shulman, and Wiesenthal (1970) observed that the experience with deception appears to make people more apprehensive of evaluation. Third, the studies by Fillenbaum (1966) and Fillenbaum and Frey (1970) caution that not all suspicious participants act upon their suspicion. Fourth, different dependent variables

seem to be differentially affected by the experience with deception. For instance, in Cook and Perrin's (1971) research, incident-learning data differed as a function of experimental history, but attitude data did not (but see Experiment 2 in Cook et al., 1970). Finally, the extent to which previous deception experience transfers to other experiments may depend on the similarity between past and present experimental situation (Brock and Becker, 1966; Cook et al., 1970).

For illustration, Page and Scheidt (1971) reported a dramatic example involving the "weapons effect," which illustrates how past experience with laboratory deception can distort behavior so extremely that it elicits a phenomenon that "cannot be generalized to nonlaboratory situations" (p. 304).[20] Page and Scheidt were able to replicate the weapons effect in only one of their three experiments, and only in a group of participants who had taken part in a deception experiment within the previous month. In contrast, participants unfamiliar with psychological experimentation never exhibited the effect. Turner and Simons (1974; see also Simons and Turner, 1976) challenged Page and Scheidt's results, and based on them Turner, Simons, Berkowitz, and Frodi (1977) came to the opposite conclusion: "Perhaps the failures to replicate the weapons effect occurred because the researchers used subjects who were not naive about deception or who were very apprehensive about the impression they might create" (p. 369). Interestingly, although Page and Scheidt (1971) and Turner et al. (1977) disagreed over the issue of how experience with deception alters experimental performance, they agreed that it does have this potential. Turner and Simons (1974) concluded: "Apparently, unless subjects are naive, the

[20] The "weapons effect" (originally reported by Berkowitz and LePage, 1967) suggests that the presence of weapons might stimulate aggression by classical conditioning processes resulting from learned associations between aggressive acts and weapons.

effects of important independent variables may be obscured"(p. 347).

To conclude, past experience with deception can, but not inevitably will, translate into non-cooperative behavior in future experiments. To what extent it does translate seems to depend on variables such as the similarity between previous and future experiments, and on the issue of how much evaluation-apprehensive behavior can interfere with the experimental responses (past experience with deception seems to promote evaluation apprehension).

4.3. *How participants' indirect experiences with deception translates:*

Individualized effects

We did not find any literature that explored directly the impact of participants' vicarious experiences on individual labs. Indeed, theoretically such effects are not likely to occur: It requires an individual experimenter acquiring a reputation for her or his deception experiments, a participant having heard about this experimenter's deception experiments through a textbook, campus scuttlebutt, or the like, and the coincidence that he or she participated in such an experiment. To the extent that many experimenters in psychology seem to restrict themselves to first-semester students –as we will see presently --, this scenario is not likely to happen. (It would be, of course, a different issue if experimenters were to rely on "professional participants", i.e., participants who are often used, as seems to be the case in many economics labs.)

4.4. *How participants' indirect experiences with deception translates:*

Generalized effects

> "It is believed by many undergraduates that psychologists are intentionally deceptive in most experiments."
> (Ledyard, 1995, p. 134)

Emotional and cognitive responses. The evidence in this category is somewhat limited. The

few available studies, however, indicate that vicarious experience may matter. According to Rubin and Moore's results (1971), for instance, it is not the number of deception experiments in which participants recall having taken part, but the number of psychology courses taken that is most closely related to suspicion. If so, then vicarious experience acquired during undergraduate psychology training (which participants in Sharpe et al.'s samples did not yet have) does matter. In fact, Higbee (1978) observed that students rated psychologists as being less truthful at the end of the semester than at the beginning (eta = .51), and students with at least five psychology courses rated psychologists as being less truthful than students who had no previous psychology course (eta = .43). Higbee concluded that "If psychologists expect the subjects to believe them, perhaps they should get the subjects at the beginning of the semester"(p.133; a refinement of the advice that Silverman, Shulman, and Wiesenthal, 1970, gave.)

Behavioral responses/Performance. While Epley and Huff (1998) found that deception experience has effects, they concluded that "naive participants, at least in this campus environment, . . . did not enter the lab with heightened suspicion, even though deception experiments are commonly conducted at our institution."(p. 766) Their finding contradicts Kelman's conjecture and the findings presented, for example, in Schultz (1969) and MacCoun and Kerr (1987). MacCoun and Kerr (1987) present the by now well-known case of a participant experiencing a genuine epileptic seizure and other participants' reactions to it (MacCoun and Kerr, 1987). Three of the five other participants (all of them introductory psychology students) reported that they questioned the authenticity of the attack and that they believed the seizure to be a charade perpetrated by the experimenter and the victim. MacCoun

and Kerr reported that "there were indications that prior knowledge of psychological research -- derived primarily from course work -- was related to suspicion. The most suspicious subject (...) had begun to study psychology in high school. Another subject recalled reading about the Milgram (1963) obedience studies in a psychology textbook."(p. 199) MacCoun and Kerr's subjects thus demonstrated indirect effects affecting even introductory psychology students, providing an example that confirms Kelman's (1967) conjecture and adds to the evidence that Schultz (1969) marshalled.

5. Discussion

Are the proponents of deception in experiments right when they argue that "there is little significant effect of deception on subject behavior or on later groups of experimental subjects" (Bonetti 1998a, p. 413; see also Kimmel, 1998)? Above we documented strong evidence for *direct* effects. Deception experiments do affect a significant fraction of participants emotionally and cognitively and provoke behavioral responses. Direct response also is bound to generate suspicion and second-guessing of the research design, and to affect judgment and decision making of a non-negligible number of participants, thus at the minimum inducing unnecessary noise in the data. Behavioral responses are clearly directed against those that use deception (and in this respect it is of less concern as the costs of deception may get internalized) and other experimenters (and in this respect there ought to be concern as the costs of deception get externalized). In sum, deception produces suspicion, non-cooperativeness, and other behavioral responses in a number of contexts.

To our initial surprise, we did not find strong evidence of the *indirect* effects hypothesized

by authors such as Kelman (1967) and Davis and Holt (1993). We believe that there are two reasons for that. First, indirect effects may be difficult to measure and to quantify, as one needs to compare participants with and without indirect experience. In fact, we have found few attempts to explicitly measure indirect effects (namely, Higbee 1978). The other example is the MacCoun and Kerr (1987) report of one participant's epileptic seizure during an experiment and the reactions to it by other participants with differential exposure to psychology. The basic problem is that few participants come to psychology experiments without strong expectations. Au contraire, the evidence suggests that psychology students come to experiments with expectations so strong that whatever they are told by the experimenter is likely to be overridden. This possibility was already suggested by Gallo et al. (1973) who reported that even in a condition in which

> subjects were told point blank that they were in a conformity experiment, quite a few speculated in the open-end question that the experiment was really a study of distance perception and that the material about being a conformity experiment was put in to throw them off the track! This type of response indicates that psychologists are disbelieved even in those rare situations in which they are caught telling the truth. (p. 106)

At the same time, even without information more than half of the participants in Gallo et al. (1973) expressed suspicion. The Gallo et al. (1973) results suggest strongly that psychologists are now in a situation where a non-deception treatment condition has become a counterfactual in many participants' minds: As a consequence, psychologists quite possibly no longer can test the baseline condition ("you will not be deceived") against experimental conditions that use

deception. It is quite possible that the baseline condition ("you will not be deceived") may, because of the documented wide-spread and frequent use of deception in psychology, no longer be implementable in psychology.

The second reason for the lack of evidence of indirect effects is the changes in the selection and composition of participant pools. Sieber and Saks (1989; see also Vitelli 1988) reported responses of 326 psychology departments with participant pools. They found that of 74% that reported having a participant pool, 93% recruited from introductory courses.[21] Often introductory students are required to take part in more than one study (within one semester) which makes it quite likely that participants have firsthand and/or vicarious experience. The authors also found that in most departments participation had to be considered nonvoluntary.[22] Thus participants are typically quite inexperienced, and the nonvoluntary nature of their participation, as well as the typical payment mode ("course credit," if any) makes it likely that participants resent the requirement.

We propose that it is likely that the peculiar institutional arrangements in psychology (namely, the widespread use of first-semester students), are the result of an evolutionary process driven by attempts to minimize both direct and indirect effects. Our view seems to be supported by Schultz (1969) who twenty years earlier, in his summary of human participant sources in 3

[21] This does not mean that 93% of their participants are from introductory courses, as 35% of the responding departments also recruit from other lower division courses. (Sieber and Saks, 1989, p. 1057)

[22] "Only 11 percent of departments have a subject pool that is voluntary in the strictest sense, that is, there are no penalties for nonparticipation, no grades for participation, and no alternatives to participation. ... most (departments) are not entirely in conformance with the APA ethical guidelines."(Sieber and Saks, 1989, p. 1058)

APA journals (1969, Table 1 on p. 217), found that on average less than 40 percent of human subjects were from introductory psychology courses. He also found that fewer participants were from introductory courses than from other courses. While the data are not completely comparable, it seems that during the two decades between Schultz (1969) and Sieber and Saks (1989), the percentage of participants from introductory courses has roughly doubled while that from non-student sources has been cut by more than half. In our view, the two snapshots presented in Schultz (1969) and Sieber and Saks (1989) suggest strongly that psychologists took the advice of Silverman et al. (1970) "that the practice of using the same subjects repeatedly be curtailed, and whenever administratively possible, subjects who have been deceived and debriefed be excluded from further participation "(p. 211) to heart. Why, if not to avoid having to deal with a contaminated subject pool, recruit from introductory courses only?[23]

Of course, both direct and indirect effects are moderated through factors such as the type of deception (e.g., Lindsay and Adair, 1990), the experienced (extent and type of harm, stress and discomfort experienced) consequences of deception (e.g., Michaels and Oetting, 1979), the quality of debriefing (Toy et al., 1989), and the nature of the experiment (scripts, repetitiveness, incentive pay, etc.; see Hertwig and Ortmann, *forthcoming*.)

What to make of Bonetti's findings and arguments? Even Anti-Prohibitionists concede that

[23] A possible alternative reason is that the wide-spread convention in psychology of not paying students might have contributed to this development. However, there is no evidence (that we know of) that the typical payment mode has changed in a manner that could provide a plausible explanation for the documented changes in participant pool selection and composition.

deception induces affective responses and changes participants' expectations. Their argument is that affective responses and suspicion (arguably an affective response and/or, albeit an imperfect, determinant for behavior) do not necessarily "cause, or establish the existence of, an alteration in laboratory behavior because of perceived deception."(Bonetti 1998, p. 381); he concluded that "data supportive of The Prohibitionist view are rare, and the balance of the evidence is certainly to the contrary."(Bonetti 1998, p. 384) Bonetti further argued that "there is little significant effect of deception on subject behavior or on later groups of experimental subjects."(Bonetti 1998a, 413).

Since Bonetti (1998, pp. 380 - 384) and others (e.g., Broeder, 1998; Kimmel, 1998; Korn, 1998), too, rely on evidence from psychology, it is necessary to discuss why we come to a different conclusion. We believe that Bonetti's review of the evidence on direct and indirect effects is seriously flawed, for at least three reasons. First, it is not systematic in his selection of the evidence, as the evidence that we amassed demonstrates his was not "an exhaustive survey of four decades of evidence"(Bonetti, 1998a, p. 413). Second, we believe his interpretation of the evidence is severely flawed.[24] Third, he completely misses the importance of the differences in

[24] Bonetti's two-page review of the evidence on direct effects is a case in point. Let us give two examples. Bonetti (1998, p. 382) argues that "Finney (1987) studied the behaviour of 120 psychology students undertaking Asch's line-judgement task. He found that informing subjects that they might be deceived did not influence conformity." True, Finney found that informing subjects they "might be deceived"did not influence conformity statistically. (The mean errors are 20% lower in the relevant baseline condition though; it is therefore necessary to point out that Finney had only 10 participants in each of his 12 treatment cells and he did not use performance-dependent payments.) However, Finney also speculated explicitly that "this information merely reaffirms subjects' prior belief that deception may occur in an experiment and, therefore, causes no change in their anticipation."(p. 45) Furthermore, Finney's results –more conformity when participants are being told that they may be deceived, and even more when they are being told that they will be deceived –are curious and counterintuitive; it is not clear what participants in

subject pools and the change that subject pools in psychology seem to have experienced over the past 3 decades.

Bonetti (1998, 1998a) argued that economists could realize significant gains from deception experiments. Form our point of view, his arguments have been effectively repudiated by Hey (1998) and McDaniel and Starmer (1998). It is, nevertheless, noteworthy that since then, Bardsley (2000) has provided a non-deception replication of Weimann (1994) which Bonetti (1998) paraded as his main example to illustrate the need for, and significant benefits of, deception experiments, demonstrating a viable alternative. It is easily predictable that such methodological innovation would slow down significantly if use of deception becomes an accepted practice among the econs. (The slowing down of methodological innovation has been acknowledged by psychologists as one of the major but little recognized costs of deception.)

---

his experiments really reacted to. Last but not least, recall that Finney provides evidence of strong cognitive and emotional responses on the part of deceived participants. Likewise, Bonnetti (1998, p. 382) argues that "Stang (1976) ... found less conformity by suspicious subjects, but suggested that this was a consequence of flawed design or deceptive conformity experiments rather than of any generalised or inevitable tendency for deception to taint behaviour." Indeed, Stang (1976) found less conformity by suspicious participants. However, contrary to what Bonetti suggests, Stang very clearly is concerned about the potential (direct and indirect) effects of deception, and the suspicion it is likely to create. To wit, "All of these determinants of suspiciousness seem to be unfortunate side-effects of previous deceptions, ... "(Stang 1976, p. 363)

6. Conclusion

It has been argued recently that there is little evidence in social science research that deception would lead to a loss of experimental control. Based on this assessment, experimental economists were counseled to lift their de facto prohibition against it in order to capture its allegedly significant benefits. Here we have assessed the costs of deception empirically and systematically, mostly drawing on evidence from psychology. Our findings suggests that there are significant cognitive-emotional as well as behavioral responses to the use of deception. To the extent that these responses are bound to introduce noise in the data, they reduce procedural regularity and replicability, and thus undermine advances that we believe experimental economists have reason to cherish (e.g., Hertwig and Ortmann, *forthcoming*). We see no good reason why economists would want to give up methodological advances for gains that are dubious at best, and most likely will produce additional costs such as the slowing down of methodological innovation. We conclude that the prohibition of deception is a sensible convention that economists should not abandon.

References

Adair, J. G. (1972). "Demand characteristics or conformity? Suspiciousness of deception and experimenter bias in conformity research." *Canadian Journal of Behavioral Science*, 4, 238-248.

Adair, J. G., Dushenko, T. W., and Lindsay, R. C. L. (1985). "Ethical regulation and their impact on research practice." *American Psychologist*, 40, 59-72.

Aitkenhead, M. and Dordoy, J. (1985). "What the subjects have to say." *British Journal of Social Psychology*, 24, 293-305.

Allen, D. F. (1983). "Follow-up analysis of use of forewarning and deception in psychological experiments." *Psychological Reports*, 52, 899-906.

Altemeyer, R. A. (1971). "Subject pool pollution and the postexperimental interview." *Journal of Experimental Research in Personality*, 5, 79-84.

American Psychological Association (1992). "Ethical principles of psychologists and code of conduct." *American Psychologist*, 47, 1597-1611.

Argyris, C. (1968). "Some unintended consequences of rigorous research." *Psychological Bulletin, 70*, 185-197.

Aronson, E. (1966). "Avoidance of inter-subject communication." *Psychological Reports*, 19, 238 ???

Asch, S.E. (1955). "Opinions and social pressure." *Scientific American* 11, 32 ???.

Asch, S.E. (1956). "Studies of independence and conformity: A minority of one against a unanimous majority." *Psychological Monographs* 70 (9, whole no. 416).

Ball, S. B. and Cech, P. A. (1996). "Subject pool choice and treatment effects in economic laboratory research." *Research in Experimental Economics*, 6, 239-292.

Bardsley, N. (2000). "Control without deception: Individual Behaviour in Free-Riding Experiments Revisited." Working paper (School of Economic and Social Studies, University of East Anglia).

Baumrind, D. (1964). "Some thoughts on ethics of research. After reading Milgram's 'Behavioral study of obedience.'" *American Psychologist*, 19, 421-423.

Baumrind, D. (1979). "IRBs and social science research: The costs of deception." *IRB: A Review*

of Human Subjects Research, 1, 1-4.

Baumrind, D. (1985). "Research using intentional deception: Ethical issues revisited." *American Psychologist*, 40, 165-174.

Berkowitz, L., and LePage, L. (1967). "Weapons as aggression-eliciting stimuli." *Journal of Personality and Social Psychology*, 7, 202-207.

Berg, J., Dickhaut, J., and McCabe, K. (1995). "Trust, Reciprocity, and Social History." *Games and Economic Behavior* 10, 122-142.

Bonetti, S. (1998). "Experimental economics and deception." *Journal of Economic Psychology*, 19, pp. 377-395.

Bonetti, S. (1998a). "Reply to Hey and Starmer & McDaniel." *Journal of Economic Psychology*, 19, pp. 411-414.

Brock, T. C. and Becker, L. A. (1966). "Debriefing and susceptibility to subsequent experimental manipulations." *Journal of Experimental Social Psychology*, 2, 314-323.

Broeder, A. (1998). "Deception can be acceptable." *American Psychologist*, 53, 805-806.

Chipman, A. (1966). "Conformity as a differential function of social pressure and judgment difficulty." *The Journal of Social Psychology*, 70, 299-311.

Christensen, L. (1977). "The negative subject: Myth, reality, or a prior experimental experience effect?" *Journal of Personality and Social Psychology*, 35, 392-400.

Christensen, L. (1988). Deception in psychological research: When is its use justified? *Personality and Social Psychology Bulletin*, 14, 664-675.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cook, T. D., Bean, J. R., Calder, B. J., Frey, R., Krovetz, M. L., and Reisman, S. R. (1970). "Demand characteristics and three conceptions of the frequently deceived subject." *Journal of Personality and Social Psychology*, 14, 185-194.

Cook, T. D. and Perrin, B. F. (1971). "The effects of suspiciousness of deception and the perceived legitimacy of deception on task performance in an attitude change experiment." *Journal of Personality*, 39, 204-224.

Coulter, X. (1986). "Academic value of research participation by undergraduates." *American Psychologist*, 41, 317.

Davis, D. D. & Holt, C. A. (1993). *Experimental economics.* Princeton: Princeton University Press.

Dickhaut, J., Hubbard, J., and McCabe, K. (1995). "Trust, reciprocity, and interpersonal history: Fool me once, shame on you, fool me twice, shame on me."(Working paper). Minneapolis: University of Minnesota, Department of Accounting.

Endler, N. S. and Hartley, S. (1973). "Relative competence, reinforcement and conformity." *European Journal of Social Psychology*, 3, 63-72.

Endler, N. S., Wiesenthal, D. L., and Geller, S. H. (1972). The generalization of the effects of agreement and correctness on relative competence mediating conformity. Canadian Journal of Behavioral Science, 4, 322-329.

Epley, N. and C. Huff (1998). "Suspicion, Affective Response, and Educational Benefit as a Result of Deception in Psychology Research." *Personality and Social Psychology Bulletin*, 24, 759-768.

Epstein, Y. M., Suedfeld, P. and Silverstein, S. J. (1973). "The experimental contract: Subjects' expectations of and reactions to some behaviors of experimenters." *American Psychologist*, 28, 212-221.

Ettinger, R. F., Marino, C. J., Endler, N. S., Geller, S. H. and Natziuk, T. (1971). "The effects of agreement and correctness on relative competence and conformity." *Journal of Personality and Social Psychology*, 19, 204-212.

Fillenbaum, S. (1966). Prior deception and subsequent experimental performance: The "faithful" subject. *Journal of Personality and Social Psychology*, 4, 532-537.

Fillenbaum, S. and Frey, R. (1970). "More on the 'faithful' behavior of suspicious subjects." *Journal of Personality*, 38, 43-51.

Finney, P. D. (1987). "When consent information refers to risk and deception: Implications for social research." *Journal of Social Behavior and Personality*, 2, 37-48.

Fisher, C. B., and Fyrberg, D. (1994). "Participant partners: College students weigh the costs and benefits of deceptive research." *American Psychologist*, 49, 417-427.

Gallo, P. S. Jr., Smith, S. and Mumford, S. (1973). "Effects of deceiving subjects upon experimental results." *The Journal of Social Psychology*, 89, 99-107.

Geller, S. H. and Endler, N. S. (1973). "The effects of subject roles, demand characteristics, and suspicion on conformity." *Canadian Journal of Behavioral Science*, 5, 46-54.

Geller, S. H., Endler, N. S. and Wiesenthal, D. L. (1973). "Conformity as a function of task generalization and relative competence." *European Journal of Social Psychology*, 3, 53-62.

Gerdes, E. P. (1987). "College students' reactions to social psychological experiments involving deception." *Journal of Social Psychology,* 107, 99-110.

Glinski, R. J., Glinski, B. C. and Slatin, G. T. (1970). "Nonnaivety contamination in conformity experiments: Sources, effects, and implications for control." *Journal of Personality and Social Psychology*, 16, 478-485.

Golding, S. L. and Lichtenstein, E. (1970). "Confession of awareness and prior knowledge of deception as a function of interview set and approval motivation." *Journal of Personality and Social Psychology,* 14, 213-223.

Gross, A. E. and Fleming, I. (1982). "Twenty years of deception in social psychology." *Personality and Social Psychology Bulletin*, 8, 402-408.

Hausken, K. (1995). "Intra-level and Inter-level Interaction."*Rationality and Society* 7, 465-488.

Hausken (1995a). "The Dynamics of Within-Group and Between-Group Interaction."*Journal of Mathematical Economics* 24, 655-687.

Hertwig, R. and Ortmann, A. (2000). "Deception and Experimental Control."Manuscript (Max Planck Institute for Human Development, Berlin.)

Hertwig, R. and Ortmann, A. (*forthcoming*). "Experimental Practices in Economics: A Methodological Challenge for Psychologists?"[Target article] *Behavioral and Brain Sciences*, forthcoming.

Hey, J. D. (1998). "Experimental economics and deception." *Journal of Economic Psychology*, 19, 397-401.

Higbee, K. L. (1978). "How credible are psychological researchers to college students? *Journal of Psychology,* 99, 129-133.

Hoffman, E., McCabe, K., and V. Smith (1996). "Social Distance and Other-Regarding Behavior in Dictator Games." *American Economic Review* 86, 653-660.

Kelman, H. C. (1967). "Human use of human subjects: The problem of deception in social psychology." *Psychological Bulletin*, 67, 1-11.

Kimmel, A. J. (1998). "In defense of deception." *American Psychologist*, 53, 803-805.

Koehler, J.J. (1996). "The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges."*Behavioral and Brain Sciences* 19, 1-53.

Korn, J. H. (1998). "The reality of deception." *American Psychologist*, 53, 805.

Kreps, D.M. (1990). *Game Theory and Economic Modelling*. Oxford: Clarendon Press.

Krupat, E. and Garonzik, R. (1994). "Subjects' expectations and the search for alternatives to deception in social psychology." *British Journal of Social Psychology*, 33, 211-222.

Landon, S. and Smith, C.E.(1998). "Quality Expectations, Reputation, and Price." *Southern Economic Journal*, 64, 628-647.

Ledyard, J. O. (1995). "Public goods: A survey of experimental research." In J. Kagel and A. E. Roth (eds.), *Handbook of experimental economics* (111-194). Princeton: Princeton University Press.

Levy, L. (1967). Awareness, learning and the beneficent subject as expert witness. *Journal of Personality and Social Psychology,* 6, 363-370.

Lichtenstein, E. (1970). "'Please don't talk to anyone about this experiment': Disclosure of deception by debriefed subjects." *Psychological Reports, 26,* 485-486.

Lindsay, R. C. L. and Adair, J. G. (1990). "Do ethically recommended research procedures influence the perceived ethicality of social psychological research?" *Canadian Journal of Behavioral Science,* 22, 282-294.

Lipton, J.P. and Garza, R.T. (1978). "Further Evidence for subject pool contamination." *European Journal of Social Psychology,* 8, 535-539.

Locksley, A., Hepburn, C. and Ortiz, V. (1982). "Social Stereotypes and Judgements of Individuals: An instance of the base rate fallacy."*Journal of Experimental Social Psychology* 18, 23-42.

MacCoun, R. J., and Kerr, N. L. (1987). "Suspicion in the psychological laboratory: Kelman's prophecy revisited." *American Psychologist*, 42, 199.

McDaniel, T. and Starmer, C. (1998). "Experimental economics and deception: A comment." *Journal of Economic Psychology*, 19, 403-409.

Mellers, Hertwig, and Kahneman (2000).

Menges, R. J. (1973). "Openness and honesty vs. coercion and deception in psychological research." *American Psychologist*, 28, 1030-1034.

Michaels, T. F., & Oetting, E. R. (1979). "The informed consent dilemma: An empirical approach." *The Journal of Social Psychology, 109,* 223-230.

Milgram, S. (1963). "Behavioral study of obedience." *Journal of Abnormal and Social Psychology* 67, 371-378.

Milgram, S. (1964). "Issues in the study of obedience: A reply to Baumrind." *American Psychologist*, 19, 848-852.

Newberry, B. H. (1973). "Truth telling in subjects with information about experiments: Who is being deceived?" *Journal of Personality and Social Psychology*, 25, 369-374.

Nicks, S. D., Korn, J. H., and Mainieri, T. (1997). "The rise and fall of deception in social psychology and personality research, 1921 to 1994." *Ethics and Behavior,* 7, 69-77.

Oliansky, A. (1991). "A confederate's perspective on deception." *Ethics & Behavior*, 1, 253-258.

Orne, M. T. (1962). "On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications." *American Psychologist*, 17, 776-783.

Ortmann, A. and Colander, D. (1997). "A Simple Principal-Agent Experiment for the Classroom." *Economic Inquiry*, 35, 443-450.

Ortmann, A., Fitzgerald, J. and Boeing, C. (*forthcoming*). "Trust, Reciprocity, and Social History: A Re-examination." *Experimental Economics.*

Ortmann, A. and Hertwig, R. (1997). "Is deception acceptable?" *American Psychologist*, 52, 746-747.

Ortmann, A. and Hertwig, R. (1998). "The question remains: Is deception acceptable?" *American Psychologist*, 53, 806-807.

Ortmann, A. and Tichy, L. (1999). "Understanding Gender Effects in the Laboratory: Evidence from Prisoner's Dilemma Games." *Journal of Economic Behavior and Organization*, 39, 327 - 339.

Page, M. M. and Scheidt, R. H. (1971). "The elusive weapons effect: Demand awareness, evaluation apprehension, and slightly sophisticated subjects." *Journal of Personality and Social Psychology*, 20, 304-318.

Rosnow, R.L. and Rosenthal, R. (1991).

Rosnow, R. L. and Rosenthal, R. (1997). *People studying people: Artifacts and ethics in behavioral research*. New York: Freeman.

Rubin, Z. (1985). "Deceiving ourselves about deception: A comment on Smith and Richardson's "Amelioration of deception and harm in psychological research." *Journal of Personality and Social Psychology*, 48, 252-253.

Rubin, Z. and Moore, J. C. Jr. (1971). "Assessment of subjects' suspicions." *Journal of Personality and Social Psychology*, 17, 163-170.

Sagarin, B.J., Rhoads, K. v. L., and Cialdini, R.B. (1998). "Deceiver's Distrust: Denigration as a a Consequence of Undiscovered Deception."*Personality and Social Psychology Bulletin* 24, 1167-1176.

Schultz, D. P. (1969). "The human subject in psychological research." *Psychological Bulletin*, 72, 214-228.

Sharpe, D., Adair, J. G. and Roese, N. J. (1992). "Twenty years of deception research: A decline in subjects' trust?"*Personality and Social Psychology Bulletin*, 18, 585-590.

Sieber, J. E., Iannuzzo, R., and Rodriguez, B. (1995). "Deception methods in psychology: Have they changed in 23 years?" *Ethics and Behavior*, 5, 67-85.

Sieber, J. E. and Saks, M. J. (1989). "A census of subject pool characteristics and policies." *American Psychologist*, 44, 1053-1061.

Silverman, I., Shulman, A. D., and Wiesenthal, D. L. (1970). "Effects of deceiving and debriefing psychological subjects on performance in later experiments." *Journal of Personality and Social Psychology*, 14, 203-212.

Simons, L. S., & Turner, C. W. (1976). "Evaluation apprehension, hypothesis awareness, and the weapons effect." *Aggressive Behavior*, 2, 77-87.

Smith, S. S. and Richardson, D. (1983). "Amelioration of deception and harm in psychological research: The important role of debriefing." *Journal of Personality and Social Psychology*, 44, 1075-1082.

Smith, S. S. and Richardson, D. (1985). "On deceiving ourselves about deception: A reply to Rubin." *Journal of Personality and Social Psychology*, 48, 254-255.

Smith, V.L. (1982). "Microeconomic systems as an experimental science." *American Economic Review* 72, 923-955.

Spinner, B., Adair, J. G. and Barnes, G. E. (1977). "A reexamination of the faithful subject role." *Journal of Experimental Social Psychology*, 13, 543-551.

Stang, D. J. (1976). "Ineffective deception in conformity research: Some causes and consequences." *European Journal of Social Psychology*, 6, 353-367.

Straits, B. C., Wuebben, P. L., and Majka, T. J. (1972). "Influences on subjects' perceptions of experimental research situations." *Sociometry*, 35, 499-518.

Stricker, L. (1967). "The true deceiver." *Psychological Bulletin*, 68, 13-20.

Stricker, L. J., Messick, S., and Jackson, D. N. (1969). "Evaluating deception in psychological research." *Psychological Bulletin*, 71, 343-351.

Taylor, K.M. and J.A. Shepperd (1996). "Probing suspicion among participants in deception research." *American Psychologist*, 51, 886-887.

Tirole, J. (1996). "A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies* 63, 1 - 22.

Toy, D., Olsen, J., and Wright, L. (1989). "Effects of debriefing in marketing research involving "mild" deceptions. *Psychology & Marketing*, 6, 69-85.

Turner, C. W., Simons, L. S., Berkowitz, L., and Frodi, A. (1977). "The Stimulating and inhibiting effects of weapons on aggressive behavior." *Aggressive Behavior*, 3, 355-378.

Turner, C. W., and Simons, L. S. (1974). "Effects of subject sophistication and evaluation apprehension on aggressive responses to weapons." *Journal of Personality and Social Psychology*, 30, 341-348.

Vitelli, R. (1988). "The crisis issue assessed: An empirical analysis." *Basic and Applied Social Psychology*, 9, 301-309.

Weimann, J. (1984). "Individual Behavior in a free-riding experiment." *Journal of Public Economics* 54, 185-200.

Wiener, R. L. and Erker, P. V. (1986). "The effects of prebriefing misinformed research participants on their attributions of responsibility." *Journal of Psychology,* 120, 397-410.

Wiesenthal, D. L., Endler, N. S. and Geller, S. H. (1973). "Effects of prior group agreement and task correctness on relative competence mediating conformity." *European Journal of Social Psychology,* 3, 193-203.

Willis, R., and Willis, Y. (1970). "Role playing versus deception: An experimental comparison." *Journal of Personality and Social Psychology,* 16, 472-477.

Wuebben, P. L. (1967). "Honesty of subjects and birth order." *Journal of Personality and Social Psychology,* 5, 350-352.

Figure 1

			Experimenter	
		Individual Reputation		Collective Reputation
	Direct effets	Contingency 1		Contingency 2
Participants				
	Indirect effects	Contingency 3		Contingency 4

The rows denote whether participants have direct experience or not; the columns denote whether experimenters deception experiments affect their own reputation only or that of the profession in general. The cells in the upper row therefore describe the repercussions of participants 'direct experiences; those in the second row describe the repercussions of indirect experiences, if any. The cells in the left column describe whether individual experimenters will be affected; those in the right column describe those repercussions for experimenters as a group.