

# **The Costs of Deception: Evidence From Psychology**

*by*

**Andreas Ortmann<sup>a</sup> and Ralph Hertwig<sup>b,\*</sup>**

<sup>a</sup>Center for Economic Research and Graduate Education, Charles University  
Economics Institute, Academy of Sciences of the Czech Republic  
Prague, CZ

<sup>b</sup>Psychology Department  
Columbia University, New York, USA

\*Correspondence: Andreas Ortmann, Ph.D.  
<mailto:ortmann@mpib-berlin.mpg.de> <mailto:ortmann@mpib-berlin.mpg.de> <mailto:aortmann@yahoo.com> CERGE-EI,  
Charles University and Academy of Sciences of the Czech Republic  
Politických veznu 7, 111 21 Prague 1, CZECH REPUBLIC  
[andreas.ortmann@cerge.cuni.cz](mailto:andreas.ortmann@cerge.cuni.cz) and [aortmann@yahoo.com](mailto:aortmann@yahoo.com)  
Tel: (011)(420)(2) 240 05 117 (office)  
*Alternatively:*  
(011)(49)(173) 65 05 191 (mobile)

Authors' note: We would like to thank Andrew Austin, Valerie M. Chase, Laura Mentz, Gerd Gigerenzer, Anita Todd, Peter M. Todd, Tom Wallsten, David Weiss, participants at the ESA/Public Choice meetings 2000 in Charleston, S.C., USA, and two referees for *Experimental Economics* for many constructive comments, and the Deutsche Forschungsgemeinschaft for its financial support (to the second author, research grant He 2768/6-1).

## Abstract

Recently, it has been argued that the evidence in social science research suggests that deceiving subjects in an experiment does not lead 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 potential benefits. To the extent that this recommendation is derived from empirical studies, we argue that it draws on a selective sample of the available evidence. Building on a systematic review of relevant research in psychology, we present two major results: First, the evidence suggests that the experience of having been deceived generates suspicion which in turn is likely to affect judgment and decision making of a non-negligible number of participants. Second, we find little evidence for reputational spillover effects that have been hypothesized by a number of authors in psychology and economics (e.g., Kelman, 1967; Davis and Holt, 1993). Based on a discussion of the methodological costs and benefits of deception, we conclude that experimental economists' prohibition of deception is a sensible convention that economists should not abandon.

## Abstrakt

V odborné literatuře zabývající se výzkumem v humanitních vědách se nedávno objevilo tvrzení, že „podvede-li“ experimentátor účastníky experimentu (tj. zatají-li před nimi pravý účel, resp. některé informace o experimentu), nemusí to nutně vést ke ztrátě kontroly nad experimentem. V důsledku toho bylo experimentálním ekonomům doporučováno, aby upustili od faktického zákazu podvodů v ekonomických experimentech, což by jim mělo umožnit využít potenciální výhody tohoto způsobu vedení experimentů. Ačkoliv se tato doporučení do jisté míry opírají o empirické studie, důkazy v jejich prospěch, jak ukazujeme, přehlíží některá dostupná fakta. Systematický průzkum relevantní psychologické literatury vede ke dvěma závěrům. Za prvé, existují důkazy o tom, že nezanedbatelné procento účastníků experimentu, kteří byli podvedeni, se stane podezřívavými vůči dalším experimentům, a to často ovlivní jejich úsudek a rozhodování. Za druhé, je jen málo důkazů pro to, že by podezřívavost vůči experimentátorovi u účastníka experimentu, který byl podveden, mohla ovlivnit i další subjekty. To je v rozporu s předpoklady mnoha výzkumníků v psychologii i ekonomii (např. Kelman, 1967 nebo Davis a Holt, 1993). Zvážení metodologických výhod a nevýhod podvodů v ekonomických experimentech nás nakonec vede k závěru, že zákaz podvodů je rozumnou konvencí, od které by ekonomové neměli upustit.

*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.<sup>i</sup> 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.

In response, experimental economists who advocate the use of deception argue that there is little evidence in the literature that deception leads to a loss of experimental control. Hence, they argue that economists' long-standing de facto proscription of deception is unduly restrictive and prevents them from capturing the potentially significant benefits of using it (e.g., and prominently, Bonetti, 1998, 1998a). 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.

Although the de facto prohibition of deception in experimental economics makes such an evaluation impossible, we can draw on data from psychology where in some areas deception is frequently used (Hertwig and Ortmann, 2001a,c). While the research agendas, practices, and subject pools of psychologists and economists diverge significantly (Hertwig and Ortmann, 2001a,b), experimenters in both disciplines share areas of common interest such as decision making in an individual and social context, thus inviting a closer look "across the border".

If the use of deception exacts a price, should it be paid? In psychology, answers to this question have been given from a consequentialist or a deontological point of view. Those favoring the latter 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 potential benefits. In contrast, those favoring the consequentialist view have argued that the costs of violating subjects' human rights need to be weighted against the costs of not being able to test theories 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<sup>ii</sup>) and that we pursue here by focusing on the methodological costs of deception. This does not mean that we take sides in the debate between consequentialists and deontologists, or that we consider the deontological arguments to be irrelevant. Rather, our view is that agreement on the merits of deception is more likely to be reached over those methodological costs of deception that can be subjected to empirical investigations.

The paper is organized as follows: In Section Two we provide a definition of deception and sketch the extent of its use in areas such as social psychology. In Section Three we discuss briefly the structure of the social situation "experiment." Building on this analysis, we review in Section Four the consequences of the use of deception. In Section Five we discuss further costs and benefits of deception, and conclude with a policy recommendation.

## 2. How frequent is the use of deception in (social) psychology?

Before we answer this question, a clarification of what constitutes deception is in order. Let us start with what is not deception. It seems undisputed among both psychologists and economists that it is not feasible, nor 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). 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).

Psychologists and economists also 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). Similarly, Nicks, Korn, and 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. Absence of full disclosure and “economy with the truth” can violate default assumptions and therefore mislead participants. We do not know what effects the violations

of default assumptions have. However, we propose that these violations, whether resultant from a sin of omission or a sin of commission, can generate suspicion and induce second-guessing of the research design. Hence violations of default assumptions should be avoided.<sup>iii</sup>

How frequent and widespread is the use of deception in psychology? Although deception is used across a wide range of areas of study such as personality, marketing, and consumer research, it appears to have been used most frequently in social 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.<sup>iv</sup>

Between 1921 and 1948, a period for which Nicks, Korn, and Mainieri (1997) analyzed the percentage of deception studies<sup>v</sup> on a yearly basis, an average of 5% of the articles in JASP employed deception each year. According to Nicks et al., the percentage rose steadily from 9% in 1948 to 51% in 1968, peaked in the 1970s (with 69% in 1975 according to Gross & Fleming, 1982), remained high in the early 1980s (with 50% in 1983; Adair et al., 1985), and declined to 31% in 1994. In an analysis of the same journal, Sieber et al. (1995) found that the percentage of deception studies dropped to 32% in 1986 but was back up to 47% in 1992. Continuing this analysis, Epley and Huff (1998) reported 42% in 1996.

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 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 APA ethics guidelines and the fact that deception studies must be reviewed by institutional review boards, deception has not become the last-resort strategy that it is meant to be.<sup>vi</sup> As we have argued elsewhere (Ortmann and Hertwig, 1997, 1998), the reasons for this may have to do with the fact that psychologists – for all practical purposes 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.

### **I. The structure of the social situation “experiment”**

To structure our presentation of the empirical evidence, it is useful first to 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 a complex game which is played inter-group (between experimentalists and participants) and intra-group (between participants or between experimentalists), and which is complicated through a power hierarchy.<sup>vii</sup> We focus below on the inter-group aspects of this game.

In the interaction between experimenters and participants, 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 to deceive participants (the principals) or to be truthful about the setting and purpose of the experiment. The principal (participant), in turn, can choose to trust the experimenter’s claims (e.g., instructions) or to doubt them. The game-theoretic predictions for this game are, depending 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 experiment's alleged purpose and nature of the experiment (at least with some probability).

The upshot is that even without previous deception experience the incentive structure of the strategic situation is such that participants who enter an experiment are likely to be suspicious. This default suspicion may be become amplified -- either through direct or indirect experience -- because of a mechanism that Davis and Holt (1993, p. 24) suggested:

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.

Talking to people who directly experienced deception is not the only way to “indirectly” experience it. Such experience can also occur through news reports about spectacular deception studies (e.g., the International Herald Tribune, September 18, 1996, on a study of the “The Bellicose Southern Male”<sup>viii</sup>), scientific articles documenting deception and participants’ reactions to it (e.g., Oliansky, 1991), and – possibly most important – undergraduate teaching of classic deception studies (e.g., Milgram’s obedience-to-authority studies). Such indirect effects are important because they transform seemingly one-shot principal-agent games into indefinitely repeated games which may have different equilibria.

In light of participants’ default suspicion, participants’ expectations that they will not be deceived (i.e., honesty on the part of the experimenter) becomes an important common good. If the experimenter chooses to deceive participants, and is found out or suspected of employing deceptive practices, then this common good might get depleted. The extent of the depletion depends on participants’ responses to firsthand (direct) or vicarious (indirect) experiences with deception: Do they choose to respond in a non-cooperative (e.g., hostile) manner and do their responses generalize beyond the individual experimenter, or her or his lab? If it generalizes, uncooperative behavior will be directed against the community of experimenters, and thus deception creates externalities.



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."

Our account of the strategic interaction between participants and experimenters suggests that any discussion of deception needs to distinguish four classes of consequences: the effects of direct experiences (i.e., firsthand experiences with deception) and indirect experiences (i.e., learning about the experiences with deception of others) on both individual experimenters and on the profession. In the following section, we discuss evidence from psychology that allows us to assess the severity of these effects. Throughout we shall distinguish whether deception affects emotion and cognition (i.e., things such as feelings, attitudes, beliefs, and expectations) and experimental performance. Of course, these categories are not mutually exclusive; indeed, some people argue that emotions are behavioral responses, or at least immediately translate into such. We will show below that evidence regarding the consequences of deception experiments is mixed. In the concluding section we consider possible reasons for the mixed picture.

#### **4. Consequences of the use of deception: Evidence**

##### *4.1. What are the effects of direct experiences with deception?*

*Emotional and cognitive responses.* Several authors have concluded that during or after the conclusion of deception experiments, participants display negative emotions. Fisher and Fyrberg (1994), among others, reported that the majority of their students believed that participants in various published deception studies must have felt embarrassed, sad, or uncomfortable than their non-deceived counterparts. Studying conformity behavior, Finney (1987) found that deceived participants believed less in the value of the research, reported less confidence in their judgment, and more tension during the experiment.<sup>ix</sup> 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 acting as confederates exhibited severe negative emotions<sup>x</sup>.

In 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)<sup>xi</sup> Interestingly, Finney (1987) also observed that those who were deceived in an experiment reported less boredom with it than those who were not deceived (despite the fact that the former experienced other negative emotions. Noting that the records of participants' reactions to deception often draw on surveys or people who imagine taking part in experiments and then report their imagined reactions, Aitkenhead and Dordoy (1985) designed an experiment that crossed active and passive (role-playing) participants and

examined the effects of minor forms of deception<sup>xii</sup>. They reported 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) .

The seemingly conflicting evidence prompts the following five observations and comments: First, the fact that deceived participants experienced less boredom in psychological experiments that suffer from a “mundane, repetitive nature” (Sharpe, Adair, and Roese, 1992, p. 589; see also Coulter, 1986) cannot justify the use of deception. Rather, such a state of affairs 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, 2001a). Second, deception (or the suspicion thereof) may in fact make a tedious and boring study less so by triggering curiosity. However, such responses -- regardless of whether they are affectively negative or positive -- undermine economists’ saliency tenet (Smith, 1982), especially if the pay-off function is flat. In other words, even if one accepted Christensen’s (1988) conclusion that participants enjoy deception experiments and believe to benefit from them, one would still face the danger that such emotional and cognitive responses add noise to the experimental data-gathering process.

Third, the findings regarding participants’ apparently positive feelings and attitudes are not as unequivocal as they were sometimes presented. For an illustration, consider the Smith and Richardson (1983) study, prominently featured by Christensen (1988, p. 668). They 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, however, the same authors 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.

Fourth, deception as used in Aitkenhead and Dordoy (1985) is not deception as used in Asch (1956) or Finney (1987) is not deception as used in Milgram (1963, 1964) or Oliansky (1991). In other words, whether deception lies within a participant’s “comfort zone” (Gerdes, 1987), and is considered acceptable by a participant, is a function of issues such as the nature and severity of deception, the methods of debriefing, and the recruitment mode. 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). Before we turn to the behavioral consequences of such suspicion, we first examine how prevalent it is among participants in psychology experiments.

*The prevalence of suspicion.* Do participants 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. This database covers the academic literature in psychology from 1887 to July 1999 (when our searches were conducted). Across the different searches that we will report below, we hunted for specific keywords in titles and abstracts. In addition, we included all psychology 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 studies cited in the articles found and included them if they seemed relevant.

In the first search, we entered the keyword “deception” in combination with “suspicion” (and its variants such as “suspicious”, “suspiciousness”, “suspiciousness”). 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, Marino, Endler, Geller, and Natziuk (1971), Endler and Hartley (1973), Geller and Endler (1973), Geller, Endler, and Wiesenhal (1973), Wiesenhal, Endler, and Geller (1973), Willis and Willis (1970), Rubin and Moore (1971), Adair (1972), and Stang (1976) found that typically between one third and two thirds of the participants had suspicions.<sup>xiii</sup>

In order to quantify the prevalence of suspicion, however, one not only has to deal with the problem of operationalizing suspicion but also with participants not admitting to be suspicious. Indeed, 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 having 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. Regardless of the experimenter's explicit instructions to the participants that they not communicate while he left the room on a pretext, Taylor and Shepperd found that participants did communicate with each other and found out that deception was involved in the experiment -- a discovery 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. 887).

Stang (1976) deserves a special mention here. After an exhaustive search of the conformity literature, he identified 21 studies that reported the percentage of participants who were classified as suspicious.<sup>xiv</sup> 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, increase in participant sophistication and decrease in participant trust.<sup>xv</sup> Specifically, Figure 1 in Stang (1976, p. 354) suggests that until the mid-sixties 1 out of 10 participants reported being suspicious, with that number shooting up to between 40 and 60 percent for the period 1967 to 1973, the last year reported in Figure 1.<sup>xvi</sup> Recall that, for example, in JPSP and its predecessor, JASP, the percentage of deception studies tripled from 16% to 47% in 1971, with similar upward swings also documented for other journals (Adair et al., 1985).

Stang (1976, p. 360) claimed that participants in conformity experiments who report suspicions commonly get excluded from the data analysis.<sup>xvii</sup> For instance, in Stang's own 1976 study, 20% of participants were excluded based on self-reported suspicions. The problem is, as we demonstrated above, it is quite possible that this number systematically

underestimates the true degree of suspiciousness since participants often may not reveal their suspicions or knowledge truthfully (e.g., Newberry, 1973; Oliansky, 1991; Taylor and Shepperd, 1996; see also Argyris, 1968; Schultz 1969). Does this non-cooperative behavior also manifest itself in participants' experimental performance?

*Experimental performance.* In general, there are two strategies to analyze 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. Alternatively, experimenters can record suspicion after participants concluded the experimental task. In what follows we explore both strategies.

*Experimentally induced suspicion.* In order to find studies that induced suspicion, we conducted another search following the procedure outlined earlier. Specifically, we used the search term "deception" in combination with "prebriefing" or "forewarning". We found eight studies (Levy 1967, Golding and Lichtenstein 1970, Gallo, Smith, and Mumford 1973, Turner and Simons 1974, Spinner, Adair, and Barnes 1977, Allen 1983, Wiener and Erker 1986, Finney 1987). Participants' knowledge and corresponding degree of 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 studies' findings, we calculated effect size measures where possible.<sup>xviii</sup> The results, once again, were mixed. For some of the eight studies we found small or no effects (e.g., Gallo et al. 1973, Allen 1983) and for some we found medium and large effects (e.g., Levy 1967, Turner and Simmons 1974, Spinner et al. 1977).<sup>xix</sup> Despite this mixed picture, 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 across a wide range of dependent variables measured in social-psychological research. In contrast, when participants were merely informed that some kind of 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 who had not been given this information (but see Spinner, Adair, and Barnes, 1977).

There are two interpretations of these results. 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 prior beliefs, it is not surprising that the experimental and the control groups (who share the same priors) do not differ.

*Suspicion recorded after participants performed the experimental task.* An alternative, though as previously described imperfect way to assess participants’ suspicion is to ask them after the experimental task (but before the debriefing) whether they had any suspicions (e.g., Asch 1956). Our search turned up fourteen studies that investigated behavior as a function of suspicion. All were concerned with conformity behavior. In ten studies, suspicious participants conformed less than unsuspecting participants. For nine of these ten studies (those in which the necessary information was given) we could calculate effect size measures (  $\eta$  or biserial correlation  $r$ ; see footnote 18): the reduction in



conformity due to suspicion was of medium to large effect size (in increasing order of effect, Adair, 1972; Stang, 1976; Endler and Hartley, 1973; Ettinger, Marino, Endler, Geller, and Natziuk, 1971; Geller, Endler, and Wiesenenthal, 1973; Rubin and Moore, 1971; Stricker, Messick, and Jackson, 1967; Geller and Endler, 1973; Glinski, Glinski, and Slatin, 1970). In four of the fourteen studies, suspicion did not significantly change the amount of conformity behavior (Chipman, 1966; Willis and Willis, 1970; Endler, Wiesenenthal, and Geller, 1972; Wiesenenthal et al., 1973). No study reported that suspicion produced greater conformity.

To conclude, the systematic samples that we investigated suggest that suspicion has the potential of altering participants' behavior in psychology and, for that matter, economics experiments. Take conformity experiments as an example: To the extent that conformity experiments explore people's compliance with social pressure, 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), using deception clearly has the potential to inject systematic error variance in the experimental data.

#### *4.2. Do participants' direct experiences with deception studies spill over to other experiments?*

*Emotional and cognitive responses.* In order to systematically assess the impact of the use of deception on participants' generalized expectations (e.g., regarding the anticipated truthfulness of information provided), we conducted yet another literature search. This time we searched for the keyword "deception" in combination with "expectation(s)" in all titles and abstracts and found – quite to our surprise – only five relevant studies. Among these studies, the size of the effect of direct experience with deception on people's generalized expectations ranged from small (Smith and Richardson, 1983; Sharpe, Adair, and Roesch 1992) to medium-large (Epley and Huff 1998, Krupat and Garonzik 1994). It is interesting to note that the most recent, and in our view most careful, studies - those by Epley and Huff and

by Krupat and Garonzik - have larger effect sizes than the earlier investigations. In what follows, we briefly describe these two studies in more detail.

Epley and Huff (1998) studied experimentally how deception impacts the suspicion of participants. 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 which explained the deceptive nature of the experiment or partial debriefing which omitted the deception part. Response to the experiment was measured through a questionnaire immediately after the debriefing. Epley and Huff found that suspicion of experiments and suspicion of the whole profession are not significantly correlated. Specifically, 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.”). Not surprisingly, participants who are made aware of the use of deception will be more suspicious in future experiments. This result, however, is important in light of the APA requirement that experimenters who employ deception must debrief participants afterwards. Clearly, this well-meaning imperative has the potential to backfire.<sup>xx</sup>

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. 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 took part involved deception, and “relatively mild forms of deception” (p. 588) to boot. In addition, psychology experiments by no means “almost always” involve deception. Thus, participants may have had the veridical 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 results and those of Sharpe et al. 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” (p. 219).

To conclude, the results suggest that firsthand experience with deception appears to increase participants’ expectations of being deceived in future experiments. This is consistent with Cook, Bean, Calder, Frey, Krovetz, and Reisman (1970, p. 189) who found that participants with a history of deception studies were more suspicious of the truthfulness of experimenters. At the same time, direct experience with deception does not seem to affect participants’ beliefs about psychologists’ trustworthiness in general.

*Experimental performance.* In order to study systematically exactly how experience with deception experiments translates into behavioral responses in future experiments, we searched the PsycINFO/PsycLIT database using the term “deception” in combination with

“experimental history”. We found a total of nine studies: Brock and Becker (1966), Fillenbaum (1966), Cook, Bean, Calder, Frey, Krovetz and Reisman (1970), Fillenbaum and Frey (1970), Silverman, Shulman and Wiesenthal (1970), Cook and Perrin (1971), Page and Scheidt (1971), Christensen (1977), and Gruder, Stumpfhauser and Wyer (1977).

In brief, the results obtained in those studies suggest 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, Silverman, Shulman, and Wiesenthal (1970) observed that experience with deception appears to make people more apprehensive of evaluation (i.e., subjects are concerned about being observed and judged, and thus will develop hypotheses about how to win positive evaluations and avoid negative ones). 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, incidental learning<sup>xxi</sup> 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 situations (Brock and Becker, 1966; Cook et al., 1970).

To illustrate how past experience with laboratory deception can distort behavior in future experiments so extremely that it elicits a phenomenon that “cannot be generalized to nonlaboratory situations” (p. 304), consider Page and Scheidt’s (1971) studies of the “weapons effect”<sup>xxii</sup>. 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 on 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 effects of important independent variables may be obscured" (p. 347).

To conclude, past experience with deception can, but does not always, 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. *Do participants' indirect experiences with deception spill over to other experiments?*

We did not find any literature that investigated the impact of participants' indirect 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 practices through a textbook, campus scuttlebutt, or the like, and the coincidence that the subject 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. Therefore, we focus here on how participants' indirect knowledge of deception studies might spill over to other experiments.

*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 matters. 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 ( $\eta = .43$ ). Based on these findings, Higbee (1978) 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.

*Behavior.* With the exception of one report, we did not find systematic attempts to investigate how indirect experience affects behavior. MacCoun and Kerr (1987) presented the case of a participant experiencing a genuine epileptic seizure and other participants' reactions to it. 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 it 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) It is interesting to note that had MacCoun and Kerr's study been concerned with altruistic behavior, then

withholding help due to suspicious would have falsely confirmed the “bystander effect” (Darley and Latané, 1968).

## **5. Discussion**

The purpose of this paper was to empirically assess the methodological costs of deception. 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)? In other words, are they right when they conclude that neither direct nor indirect experiences with deception have consequences? We strongly disagree with the first part of their conclusion. Our systematic review of the literature has yielded evidence that direct experience with deception can affect participants emotionally and cognitively. Most importantly, being suspicious of deception while participating in an experiment may powerfully alter behavioral responses (see conformity experiments). In addition, there is evidence to suggest that direct experience spills over into future experiments by eliciting the expectation of being deceived again. Based on these findings, it seems fair to conclude that direct experience with deception has the potential to alter the experimental performance of those participants who entertain suspicions, and to generate suspicion and second-guessing in future experiments.

We agree with Bonetti, however, that there is no clear evidence that indirect experience with deception matters, as authors in psychology (Kelman, 1967) and in economics (Davis and Holt, 1993) have hypothesized. We believe that there are two reasons for our finding. First, indirect experiences (e.g., via campus scuttlebutt or media exposure) are simply difficult to measure and quantify, as one needs to compare participants with and without indirect experience. This difficulty may explain why few attempts have been made to study the effect of indirect experience. If so, then the lack of such studies merely suggests

that the issue is difficult to investigate, but not that indirect experience of deception comes without methodological costs.

A second reason for the lack of evidence regarding the consequences of indirect experiences may be that they may not exist in psychology. This does not mean, however, that indirect experiences do not exact a price. Rather, psychologists try hard to avoid paying the price by relying on students from introductory courses who may still be naive. In a relatively recent review, Sieber and Saks (1989; see also Vitelli 1988) documented the extent to which psychologists depend on such students. Specifically, they 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.<sup>xxiii</sup> For comparison, consider Schultz's (1969, Table 1) summary of the composition of the human subjects pool (in studies published in three APA journals) published twenty years earlier. He reported that on average less than forty percent of human subjects were from introductory psychology courses. Albeit speculative, it is plausible that psychology's peculiar institutional arrangements (to recruit subjects mostly from introductory courses) could be the result of an evolutionary process - a process driven by attempts to minimize the contaminating effects of deception on the subject pool (and certainly also to guarantee "free" access to subjects.) If true, psychologists may have taken the advice of Silverman et al. (1970) to heart "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 conclude, we agree with previous observations that there is little evidence that indirect experience with deception has negative effects. However, we suggest that those effects might very well become manifest if recruitment practices in psychology would not curtail them, or alternatively, if a discipline such as experimental economics, which does not recruit students primarily from introductory classes as experimental subjects, began using deception.



*To deceive or not deceive?* We believe that Anti-Prohibitionists such as Bonetti (1998a) are to be applauded for reflecting on the costs and benefits of deeply entrenched methodological practices. Like Bonetti, we believe that in order to put our practices in perspective, it is useful to look at methodological practices across disciplines (Hertwig & Ortmann, 2001 a,b). Based on our own extensive and systematic review of studies on the consequences of deception conducted in psychology, however, we disagree with Bonetti's (1998a) conclusion, based on his "brief review of the available evidence" (p. 384) that "the experimental evidence directly undercuts the basis of attempts to proscribe deception" (p. 389). We believe that his conclusion is the result of a less extensive and systematic review of the evidence.

According to Bonetti (1998a) the most important benefit of deception is that "it is the way in which the attention of the subjects can be effectively distracted, thus ensuring that the behavior which is measured is more natural and spontaneous and less affected and contrived" (p. 386). In other words, Anti-Prohibitionists argue that certain aspects of human behavior can only be studied if people are caught off guard. Bonetti's argument echoes that of advocates of deception in psychological experimentation who argue that, for instance, in investigations of socially undesirable aspects of behavior the experimenter needs to camouflage the purpose of the experiment to achieve experimental control. If not, the "psychologist runs the risk of distorting the reactions of his or her subjects and ultimately limiting the applicability of the research findings" (Kimmel, 1996, p. 68). This argument, as persuasive as it may look, ignores the empirical evidence that we have accumulated above and that, to our mind clearly, indicates that the direct experience of deception has the potential to change experimental performance. In other words, the very use of deception can impair, and even destroy, the experimental control it is meant to achieve. Moreover, its use exacts another cost, which needs to be taken into account when trying to strike a balance

between the costs and benefits of deception, namely the slowing down of methodological innovation.

This argument was well expressed three decades ago by Baumrind (1971, p. 893):

Many of the investigators who choose to use Machiavellian means in experimental settings are brilliant and creative methodologists. The likelihood is that if such men knew that in order to investigate experimentally an area in which they were interested they would have to revise their research strategy, they would ... be capable of inventing new experimental methods that were well controlled as well as humane.

Indeed, given that the alleged last-resort strategy of deception is widely accepted in psychology, why bother developing an experimental methodology that would allow psychologists to do without? We argue that the innovation slowdown is an important cost that economists would incur were they to give up their de facto prohibition against deception. In this context, it is noteworthy that Bardsley (2000) recently provided a non-deception replication of Weimann (1994), which Bonetti (1998) paraded as a prime example of an economic investigation that “necessarily required deception” (p. 387). Bardsley’s innovative design demonstrates that this statement is untenable.<sup>xxiv</sup> In fact, we suggest that there is no theory in economics that could not be tested without deception.

To conclude, using evidence from psychology we have attempted to empirically and systematically assess the methodological costs of deception. The review of the available evidence suggests that the direct experience of deception and the suspicion of deception carry with them the potential of provoking significant cognitive-emotional as well as behavioral responses. To the extent that these responses are bound to introduce systematic error variance in the data, they impair, and possibly destroy, experimental control. In light of this danger and the other potential costs (e.g., added subject pool selection biases or the slowing down of methodological innovation if deception were to be accepted among economists), 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.
- 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.
- Asch, S.E. (1956). "Studies of independence and conformity: A minority of one against a unanimous majority." *Psychological Monographs* 70 (9, whole no. 416).
- Bardsley, N. (2000). "Control Without Deception: Individual Behaviour in Free-Riding Experiments Revisited." *Experimental Economics* 3, 215 - 240.
- Baumrind, D. (1964). "Some thoughts on ethics of research. After reading Milgram's 'Behavioral study of obedience.'" *American Psychologist*, 19, 421-423.
- Baumrind, D. (1971). "Principles of ethical conduct in the treatment of subjects: Reaction to the draft report of the Committee on Ethical Standards in Psychological Research." *American Psychologist*, 26, 887\_896.
- 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. (1974). Some determinants of impulsive aggression: Role of mediated associations with reinforcements for aggression. *Psychological Review*, 81, 165\_176.
- Berkowitz, L., and LePage, L. (1967). "Weapons as aggression-eliciting stimuli." *Journal of Personality and Social Psychology*, 7, 202-207.
- Bonetti, S. (1998). "Experimental economics and deception." *Journal of Economic Psychology*, 19, 377-395.

- Bonetti, S. (1998a). "Reply to Hey and Starmer and McDaniel." *Journal of Economic Psychology*, 19, 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.
- 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.
- Darley, J. and Latané, B. (1968). "Bystander intervention in emergencies. Diffusion of responsibility." *Journal of Personality and Social Psychology*, 8, 377\_383.
- Davis, D. D. and 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., Wiesenhal, 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.

- 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 Wiesenhal, 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.
- Gruder, C. L., Strumpfhauser, & Wyer, R. S. Jr. (1977). "Improvement in experimental performance as a result of debriefing about deception." *Personality & Social Psychology Bulletin*, 3, 434-437.
- 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. (2001a). "Experimental Practices in Economics: A Methodological Challenge for Psychologists?" *Behavioral and Brain Sciences* 24, 383 - 403.
- Hertwig, R. and Ortmann, A. (2001b). "Money, lies, and replicability: On the need for empirically grounded experimental practices and interdisciplinary discourse." *Behavioral and Brain Sciences* 24, 433 - 444.
- Hertwig, R. and Ortmann, A. (2001c). "Economists' and Psychologists' Experimental Practices: How They Differ, Why They Differ, And How they Could Converge," in I. Brocas and J. Carrillo (eds.), *Psychology and Economics*, Oxford University Press, 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. (1996). *Ethical issues in behavioral research: A survey*. Cambridge: Blackwell Publishers.
- Kimmel, A. J. (1998). "In defense of deception." *American Psychologist*, 53, 803-805.
- 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.
- 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.
- 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.

- 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 and Behavior*, 1, 253-258.
- Ortmann, A. and Colander, D. (1997). "A Simple Principal-Agent Experiment for the Classroom." *Economic Inquiry*, 35, 443-450.
- 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.
- Rosenthal, R. and Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis (2nd ed.)*. McGraw Hill.
- 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., and Turner, C. W. (1976). "Evaluation apprehension, hypothesis awareness, and the weapons effect." *Aggressive Behavior*, 2, 77-87.
- Smith, V.L. (1982). "Microeconomic systems as an experimental science." *American Economic Review* 72, 923-955.
- 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.
- 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.



- Toy, D., Olsen, J., and Wright, L. (1989). "Effects of debriefing in marketing research involving "mild" deceptions." *Psychology and Marketing*, 6, 69-85.
- 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.
- 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.
- Vitelli, R. (1988). "The crisis issue assessed: An empirical analysis." *Basic and Applied Social Psychology*, 9, 301-309.
- Weimann, J. (1994). "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.

---

<sup>i</sup> 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 2001a).

<sup>ii</sup> 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 suggests that deception is not perceived as such. In many areas of psychology, deception is indeed considered a n inconsequential transgression that “unrepentant deceivers” (as one psychologist called himself in a communication to one of us) rationalize with the elegance of a research design and that others (to cite an example that one of us encountered recently as a referee) are allowed to rationalize with poorly programmed experiments. In Hertwig and Ortmann (2000) -- written for a psychology audience and a companion paper of sorts -- we propose an incentive-compatible mechanism designed to reduce the frequency of deception experiments in psychology.

<sup>iii</sup> Admittedly, this statement is difficult to operationalize. In the words of one referee, “How will we know what range of default assumptions might be entertained by a given sample pool; how much time/effort should be invested to find out; how hard should we work to convince subjects to revise their default assumptions?” There are, in our view, no general answers to these good questions. Ultimately, any specific answer is a function of the particular circumstances of place and time and it has to be a judgment call of the individual experimenter. One sensible strategy of assessing default assumptions are appropriately constructed pilot sessions. In addition, a monitor (i.e., a randomly selected subject) may also serve as a source of information about participants’ typical default assumptions.

<sup>iv</sup> We found 14 studies that analyzed the frequency of deception across a wide range of journals and areas of study (e.g., Toy, Olsen, and Wright, 1989). We use JASP and JPSP for illustration because the data for these journals are the most comprehensive.

<sup>v</sup> Following the terminology in the psychology literature, below we call a study (an experiment) that involved deceptive methods a “deception study” (“deception experiment”). This does not necessarily mean that the effects of deception were studied, as the shorthand expression might suggest.

<sup>vi</sup> 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 has succeeded in reducing the severity of deceptive methods used. 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 appears to be substantial (e.g., Oliansky, 1991, or Asch, 1956).

<sup>vii</sup> The relationship between participants and experimentalists “has some of the characteristics of a 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.” (Schultz, 1969, p. 221) Similarly, Argyris (1968) argued that “[subjects] are now beginning to behave like lower level employees in companies” (p. 187). Such a power hierarchy 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. According to an analysis of recruiting practices in

---

psychology, “[o]nly 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)

<sup>viii</sup> In an attempt to compare stress and aggression related hormone surges for southern and northern white males, the authors of the study instructed confederates to bump into participants and call them insulting names. A couple of years after the study came out, one of us, while taking in the spectacular vista of the Monhegan Island coast (Maine, USA), overheard one hiker tell another about that very experiment.

<sup>ix</sup> Following Asch (1956), conformity experiments typically place a subject in a judgment or decision situation (e.g., the line-length comparison task) together with other “participants”. These other participants, however, are confederates of the experimenter and are instructed to make incorrect claims. The purpose of such an experimental scenario is to study compliance with social pressure, and adherence to social norms.

<sup>x</sup> 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.

<sup>xi</sup> Curiously, Christensen’s review did not contain his own earlier article in which he concluded 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.” (Christensen, 1977, p. 399)

<sup>xii</sup> 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 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. The Aitkenhead and Dordoy participants were not paid; Finney’s participants were offered extra credit points that were, however, mostly inconsequential.

<sup>xiii</sup> These numbers resulted from post-experiment interviews.

<sup>xiv</sup> 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).

<sup>xv</sup> Stang (1976, p. 355) mentions other possible explanations for the increase in suspicion, namely decreased quality of the cover story and measurement artifacts. Both strike us as unlikely.

<sup>xvi</sup> Stang (1976, p. 354) states that the percentage of subjects correlates .76 ( $df = 19, p > .05$ ) with the year the study was published.

---

<sup>xvii</sup> Conformity experiments, however, are by no means the only areas of research where suspicious participants (self-reported) get excluded (e.g., Sagarin, Rhoads, and Cialdini, 1998).

<sup>xviii</sup> 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 F has only a single df in the numerator, as is the case when two conditions are compared (which happened 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.

<sup>xix</sup> All effect size computations of this kind (including those reported below) and more detailed discussion of the manipulations and behavioral effects of each study reported here, may be found in Hertwig and Ortmann (2000; see especially Tables 2 - 5). See <http://home.cerge.cuni.cz/Ortmann/recentWPs.html>.

<sup>xx</sup> Toy, Olsen, and Wright (1989) reported that debriefing happens rarely in marketing research and conjectured that experimenters thus try to avoid this problem.

<sup>xxi</sup> Incidental learning is an experimental paradigm used to investigate learning without intent.

<sup>xxii</sup> The weapons effect (originally reported by Berkowitz and LePage, 1967) describes the observation that the mere sight of guns can facilitate aggressive thoughts and behavior (e.g., in terms of number of shocks that a subject inflicts on a confederate subject). Berkowitz (1974) explained this finding in terms of a classical conditioning process: Weapons are associated with aggressive stimuli through their frequent pairing with aggressive acts in real or fictional life. Thus, when an aroused and uninhibited person is exposed to a weapon, it might elicit the responses that have frequently been associated with guns, i.e., aggression\_facilitating reactions.

<sup>xxiii</sup> 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)

<sup>xxiv</sup> The subjects in Bardsley (2000) were told that some of the public good scenarios in the experiment will be "fictitious". Bardsley (2000) shows that under certain conditions it is a dominant strategy for subjects to treat these scenarios as "real". His approach allows him to study how subjects react to specific patterns of contributions by others - the major goal of Weimann (1994) who used deception to achieve this goal.