

Running head: DECEPTION IN EXPERIMENTS

Deception in Experiments:
The Costs of an Alleged Method of Last Resort

Ralph Hertwig

Max Planck Institute for Human Development, Berlin

Andreas Ortmann

Charles University and Academy of Sciences of the Czech Republic, Prague

Send correspondence to:

Ralph Hertwig

Max Planck Institute for Human Development

Center for Adaptive Behavior and Cognition

Lentzeallee 94

14195 Berlin, Germany

Tel: +49-30-82406-354

E-mail: hertwig@mpib-berlin.mpg.de

Abstract

In psychology, deception is commonly used to increase experimental control. Yet, its use has provoked concerns that it raises participants' suspicions, prompts second-guessing of experimenters' true intentions, and ultimately distorts behavior and the control it is meant to achieve. These concerns can and have been subjected to empirical analysis. Our review of the evidence yielded two key results: First, there is evidence that participants who experienced deception firsthand are likely to become suspicious and that there are non-negligible differences between suspicious and reportedly naïve participants. Second, there are surprisingly few studies addressing the question of whether suspicion can result from secondhand experience with deception such as undergraduate psychology training or the profession's reputation more generally. In light of the latter finding, we propose an incentive-compatible mechanism designed to encourage researchers to search for and implement alternatives that forego deception. Thus, making this tool truly the strategy of last resort, as intended by the rules of conduct of the American Psychological Association.

The use of deception [in experiments] has become more and more extensive.... It is easy to view this problem with alarm, but it is much more difficult to formulate an unambiguous position on the problem.... I am too well aware of the fact that there are good reasons for using deception in many experiments. There are many significant problems that probably cannot be investigated without the use of deception, at least not at the present level of development of our experimental methodology (Kelman, 1967, p. 2).

In his well-known article “Human use of human subjects: The problem of deception in social psychology,” published in the pages of Psychological Bulletin, Herbert Kelman (1967, p. 2) described his dilemma as a social scientist as that of being caught between the Scylla of the use of deception to study important social behaviors, and the Charybdis of ethical and methodological considerations. He wrote this article in the wake of a public exchange between Baumrind (1964) and Milgram (1964) and in response to the substantial increase in the use of deception during the 1960s. Whereas the exchange between Baumrind and Milgram focused on the ethical implications of Milgram’s research on obedience, Kelman (1967) stressed the long-term methodological consequences of deceptive practices on participants’ expectations and behavior. The essence of his concern is expressed in this prophetic statement:

As we continue to carry out research of this kind, our potential subjects become increasingly distrustful of us, and our future relations with them are likely to be undermined. Thus, we are confronted with the anomalous circumstance that the more research we do, the more difficult and questionable it becomes. (Kelman, 1967, p. 7)

We recently reiterated Kelman’s concern that the use of deception may contaminate the participant pool (Ortmann & Hertwig, 1997, 1998), and to put psychology’s research practices into perspective we noted that researchers in a neighboring discipline, experimental economics, have effectively prohibited the use of deception in their experiments (Hertwig & Ortmann, 2001, 2002; Ortmann & Hertwig, 2002a). Our comments prompted responses from several researchers, Bröder (1998), Kimmel (1998), Korn (1998), and Weiss (2001). Four arguments feature prominently in their defense of deception. They are: (1) The use of deception has, after an increase in the 1960s and 1970s, dropped (e.g., Korn, 1998); (2) “the preponderance of evidence suggests

that deceived participants do not become resentful about having been fooled by researchers” (Kimmel, 1998, p. 804); (3) the effects of suspiciousness on research performance “appear to be negligible” (Kimmel, 1998, p. 804); and (4) deception is an indispensable tool for achieving experimental control—at least in some socially significant areas of research (Bröder, 1998; Kimmel, 1998; Korn, 1998; Weiss, 2001).

Our original contributions (Ortmann & Hertwig, 1997, 1998), as well as those by our critics, were based more on assertions than on empirical evidence. It is thus the main goal of this paper to present empirical evidence that bears on these four arguments. Our focus will be on the possible methodological side effects of deception. We do not address the ethical issues that inextricably are linked with the use of deception. This is not to say that we believe these concerns, laid out eloquently by Baumrind (1964, 1971, 1985; see also Herrera, 1996; Aguinis & Handelsman, 1997; Kimmel & Smith, 2001), are irrelevant.¹ However, more than 30 years of controversy over the ethical issues have made it clear that this debate is necessarily driven by values, sometimes categorical values, over which people are bound to disagree. Agreement, we believe, may be more easily achieved with regard to the reality (or lack thereof) of methodological costs of deception—once the available evidence has been systematically compiled. To make progress toward a comprehensive collection of the evidence is the purpose of the present paper. We begin by providing a definition of deception and describing the reasons for its use.

What Is Deception?

Deception is not easily defined. Yet, there seems to be considerable agreement about what definitely ought to count as deception. Such agreement is, for instance, manifest among the group of researchers who have studied the prevalence of deception as a research method in (mainly social) psychology. We found a total of 14 studies that analyzed the frequency of use of deception in various journals.² Examination of their criteria for defining deception reveals that intentional and explicit misrepresentation, that is, lying about, for instance, the purpose of the investigation and the identity of researcher and confederate, is unanimously considered to be deception. This consensus is also shared across disciplinary borders. In the words of economist Hey (1998), “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” (p. 397, his emphasis).

Hey's assertion, furthermore, indicates what seems indeed to be widespread agreement among researchers: Withholding information does not necessarily constitute deception. That is, not acquainting participants in advance with all aspects of the research being conducted, such as the hypotheses explored and the full range of experimental conditions is typically not considered deception. In their review of deception studies, Adair et al. (1985), for instance, decided that "the simple failure to disclose the true purpose of the study was not counted as deception" (p. 63). Although Baumrind (1979, p. 1) suggested that "full disclosure of everything that could affect a given subject's decision to participate is a worthy idea," this strict critic of deception also conceded that "absence of full disclosure does not constitute intentional deception" (Baumrind, 1985, p. 165). Similarly, experimental economists McDaniel and Starmer (1998, p. 406) described some forms of "economy with the truth" as "perfectly legitimate," and Hey (1998, p. 397) pointed out that "ill-defined experiments" (i.e., when the experimenter does not inform participants about all features of the experiment) are an important tool (see Lawson, 2001, for a thorough discussion of the distinction between providing false information and withholding information).

The distinction between deception and non-deception blurs, however, when participants' default assumptions come into play. One default assumption a participant is likely to have is that experiments start only after an experimenter has clearly indicated its beginning. As a consequence, she might assume that her initial interactions with the experimenter (upon entering the laboratory) are not an object of investigation. Should violations of such expectations be counted as deception? Some of the researchers who assessed the prevalence of deception did not appear to include such violations (Gross & Fleming, 1982; Adair et al., 1985; Nicks et al., 1997), but others did. Sieber et al. (1995), and Gross and Fleming (1982), for instance, considered participants to be deceived if they were unaware of being research participants at all or were unaware that the study had begun at the time of the manipulation. The fact that some researchers included violations of default assumptions in their definition of deception and others did not might reflect conceptual disagreement. Alternatively it could reflect a pragmatic decision on the part of researchers who struggle to quantify the prevalence of deception—violations of default assumptions are much more difficult to identify than provisions of misinformation. We conjecture that violations of default assumptions have a similar potential for creating suspicion and distrust as does the provision of false information and they should therefore be treated as deception.

In sum, a consensus has emerged across disciplinary borders that intentional provision of misinformation is deception and that withholding information about research hypotheses, the range of experimental manipulations, or the like ought not to count as deception. Common ground has not (yet) been established with respect to the violation of participants' default assumptions.³

Reasons for Deception and Two Degrees of Contamination

Why deceive? Deception is often justified with two arguments. The first argument is that deception allows the researcher to create situations of special interest that are not likely to arise naturally. A good illustration of this strength is found in studies of helping behavior in emergency situations, in which researchers stage emergencies (e.g., someone experiences a seizure), manipulate situational factors (e.g., absence and presence of others), and then determine the impact of these factors on bystanders' willingness to help (e.g., Darley & Latané, 1968). Fortunately, emergency situations occur infrequently; this fact, however, makes them very difficult to study experimentally, unless, so the argument goes, one fabricates them.

The second rationale for deception is that certain socially relevant aspects of behavior can only be studied if people are caught off guard (e.g., Weber & Cook, 1972; Cooper, 1976; Weiss, 2001). If they suspected or knew that some socially undesirable aspects of behavior are being observed (e.g., conformity, prejudices, anti-social behavior), then they would alter their "natural" behavior to look as good as possible to the social observers (i.e., experimenter or other participants). Consider conformity behavior as an example. If participants knew that an experiment explores the extent to which they easily give in to social pressure, then they would be less likely to show conformity behavior. Therefore, so the argument goes, studies of conformity behavior need to camouflage the purpose of the experiment to achieve experimental control. If not, then "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).

Challenging the latter rationale, critics of deception have argued that it is the very use of deception that impairs, even destroys experimental control, thus threatening the validity of research findings. Kelman (1967) is not the only one to have advanced this argument. Other researchers in the social sciences, that is, psychologists, sociologists, anthropologists, and experimental economists, also have worried that deception contaminates the participant pool. While in sociology it was suggested that a likely outcome of deceptive practices is participants' future resistance to other

research efforts (e.g., Erikson, 1996), psychologists and experimental economists have expressed the concern that the expectation of being deceived produces suspicion and second-guessing, and that these reactions rather than the experimenter's scenario and instructions guide and ultimately distort experimental behavior. Illustrations of expressions of this concern are listed in Table 1.

[Insert Table 1]

These statements differ in one important aspect. Whereas some researchers attribute the contamination of the participant pool to firsthand experience with deception (i.e., participating and being debriefed in deception experiments; see Seeman, 1969), others assume that secondhand experience with deception (e.g., stemming from undergraduate psychology classes, campus scuttlebutt, media coverage of psychological research, and the profession's reputation more generally) suffices to engender in participants the expectation that they will be deceived (e.g., Orne, 1962; Ring, 1967; Adelson, 1969; Hey, 1991; Davis & Holt, 1993; Ledyard, 1995). This assumption is particularly common in experimental economics where deception is effectively prohibited (e.g., see Hey, 1991; Davis & Holt, 1993, in Table 1). In the parlance of economists, participants' expectation that they will not be deceived (i.e., that experimenters will be honest) is a common good such as air or water that would be depleted quickly even if only a few of their tribe practiced it, hence experimental economists' strict proscription of deception despite occasional calls to relax this rule (e.g., Bonetti, 1998a, b).

The distinction between firsthand and secondhand experiences is relevant because these different experiences imply different degrees of contamination of the subject pool (if there is any at all). If secondhand experience sufficed to induce suspicion and second-guessing, then the potential side effects of deception would likely be widespread and extend beyond participants with firsthand experience. In contrast, if firsthand experience were necessary to induce suspicion and second-guessing, then the degree of contamination would be more contained. In addition, the argument advanced to defend the use of deception—that its use has dropped since its peak in the 1970s—would then gain additional weight.

Argument 1: Is the Use of Deception in Decline?

As mentioned earlier, we found 14 studies (see Footnote 2) that analyzed the use of deception across a wide range of journals (deception is by no means confined to social psychology; see, for instance, Toy, Olsen, & Wright, 1989)⁴. Here we focus on the results for the highly ranked

Journal of Personality and Social Psychology (JPSP) (and its predecessor, the Journal of Abnormal and Social Psychology, JASP), for which the most comprehensive and recent figures are available.

Between 1921 and 1948, a period for which Nicks et al. (1997) analyzed the use of deception on a yearly basis, an average of 4.7% of the articles in JASP employed deception each year. According to the same authors, the percentage rose steadily from 9% in 1948 to 50.7% in 1968. It then peaked in the 1970s (with 69% in 1975 according to Gross & Fleming, 1982), remained high in the early 1980s (with 49.7% in 1983; Adair et al., 1985) and declined to 31.3% in 1994. In their 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. Some of the fluctuations may reflect substantial changes in the applied methods (e.g., the initial upswing in the 1960s), ethical standards, and federal regulation of research. Others may reflect different definitions of what constitutes deception (e.g., compare the more inclusive criteria employed by Sieber et al., 1995, with the criteria used by Nicks et al., 1997).

Although these figures show a clear decline compared to the heyday of deception in the late 1960s and 1970s, they also demonstrate that the absolute level, compared to psychology's past, is still high: A conservative estimate is that every third study published in JPSP in the 1990s employed deception compared to 4.7% between 1921 and 1948. For other social psychology journals, such as the Journal of Experimental Social Psychology (JESP), the proportion appears to be even higher (e.g., Gross & Flemming, 1982; Nicks et al., 1997), and in some applied areas such as marketing research the number of deception studies has actually risen over recent decades (from 43% in 1975-1976 to 56% in 1996-1997 for the Journal of Marketing Research and Journal of Consumer Research; see Kimmel, in press).

Returning to the distinction of firsthand and secondhand experiences with deception, these figures suggest that for students to experience deception personally is not an unlikely event—even today. Thus, in our view, arguing that its frequency has dropped cannot easily dismiss the concerns about possible methodological side effects of deception. The absolute level of the use of deception remains high (e.g., by some historical standards), and it is by no means a method of last resort. We now turn to the second argument in defense of deception. According to this argument, deceived participants do not become resentful about having been fooled by researchers.

Argument 2: Does Deception Breed No Resentment?

Based on his review of research about the effect of deception, Christensen (1988) concluded: “This review of the literature, which has documented the impact of deception on research participants, has consistently revealed that research participants do not perceive that they are harmed, and do not seem to mind being misled (p. 668).⁵ More recently, Kimmel (1998) conjectured that the “preponderance of evidence suggests that deceived participants do not become resentful about having been fooled by researchers” (p. 804).

According to the Merriam-Webster’s Collegiate Dictionary, “resentment” is “a feeling of indignant displeasure or persistent ill will at something regarded as a wrong, insult, or injury.” Does the empirical evidence in fact indicate that participants do not harbor such feelings and do not mind being misled? Kimmel’s (1998) conclusion seems to rest on the following five observations: (1) Participants, in general, do not seem to express negative feelings (e.g., regret having participated) about their experience in deception experiments (e.g., Milgram, 1964; Ring, Wallston & Corey, 1970; Smith, 1981; Pihl, Zacchia & Zeichner, 1981).⁶ (2) Participants endorse the scientific utility of deception experiments (Clark & Word, 1974; Gerdes, 1979), and seem to be prepared to tolerate deception in the interest of research (Aitkenhead & Dordoy, 1985). (3) Participants in deception experiments report having enjoyed the experience more, feeling less bored, and perceiving more educational benefit from their participation than participants in non-deception experiments (e.g., Smith & Richardson, 1983; Finney, 1987).⁷ (4) Most college students are generally accepting of ethically sensitive research practices such as deception and invasion of privacy (e.g., Epstein, Suedfeld & Silverstein, 1973; Farr & Seaver, 1975; Collins, Kuhn, & King, 1979) and are less critical of those practices than members of Human Subjects Committees, psychologists, graduate students, and faculty (e.g., Sullivan & Deiker, 1973; Smith, Berard, & Malinowski, 1980; Smith & Berard, 1982; Korn, 1987). (5) Finally, according to a questionnaire study by Sharpe and colleagues (1992), the continued use of deception has not evoked an increase in negative attitudes toward psychological research among the participant population.⁸

Based on this evidence, Kimmel (1996) concluded that “the negative effects of deception appear to be minimal” (p. 104). But are they really? A series of other observations provide less reason for such optimism. Fisher and Fyrberg (1994), for instance, reported that the majority of their students believed that participants in various published deception studies must have felt embarrassed, sad, or uncomfortable. In one experiment, Allen (1983) found that only 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, & Majka, 1972). Moreover, 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. In addition, Epstein et al. (1973) reported that, next to danger to the participant, deception is the most frequently mentioned reason to withdraw from an experiment. Oliansky (1991) observed that deception (here: having the impression that one has control over another participant, who was in reality a confederate) might trigger severe negative emotions in (some) participants.

We can think of two reasons for why the evidence regarding people’s feelings is so mixed. First, deception as used in Aitkenhead and Dordoy (1985) is not deception as used in Asch (1956) or Finney (1987), which is not deception as used in Milgram (1964) or Oliansky (1991). In other words, whether being fooled does or does not lie within a participant’s “comfort zone” (Gerdes, 1979) is probably a function of the nature and severity of the deception. Second, participants react on different levels, and their responses need not converge: For instance, in a replication of Asch’s line-judgment task, Finney (1987) observed that deceived participants were more depressed, hostile, and anxious than non-deceived participants; yet, their uneasiness did not cause them to avoid future psychological research or to question the study’s scientific value.

Whatever the reasons for the mixed results may be, it seems fair to conclude that the issue of whether deception raises resentment is not yet decided. It appears even less settled when one consults related research on the consequences of deception in social interactions. The results of a still small set of negotiation and strategic interactions (i.e., games) studies suggest that being deceived in social interactions has the potential to evoke a wide range of responses, ranging from diminishing desire for future interactions, attribution of untrustworthiness (Boles, Croson, & Murnighan, 2000) to a substantial taste for retribution and for punishment by the deceived players—a taste for which they are even willing to sacrifice money (e.g. Boles et al., 2000; Brandts & Charness 2002; Croson, Boles, & Murnighan, in press).⁹ However, the same research also suggests another reason why studies in psychology may have arrived at contradictory conclusions about people’s resentment of deception (or lack thereof). In an analysis of people’s evaluations of ethically questionable negotiation tactics, Lewicki and Stark (1996) observed that players’ expectations of the “game”

being played mitigate their responses: If people expect lies and deception, then they might not necessarily respond negatively once their expectations become reality. This finding, however, provides little comfort because it implies that those students who do not resent deception may be the ones who take deception simply as part of the game.

Argument 3: Are the Effects of Suspicion Negligible?

Psychological experiments appear to provoke a peculiar dynamic that Riecken (1962) described as follows:

The fact that the experimenter controls the information available to the subject and that he never reveals completely what he is trying to discover or how he will judge what he observes—this feature gives the experiment much of its character as a game or contest. It leads to a set of inferential and interpretive activities on the part of the subject in an effort to penetrate the experimenter's inscrutability.

Notwithstanding their possible endorsement of the use of deception, are not participants who suspect the experimenter to lie even more eager to undo the information asymmetry and, in Riecken's words, to penetrate the experimenter's inscrutability? If so, one may expect the behavior of participants in experiments who suspect foul play to differ from those who do not. Based on his review of the literature, Kimmel (1998), however, arrived at the opposite conclusion. In his view, "the effects of suspiciousness on research performance, though somewhat inconsistent, appear to be negligible, leading some to conclude that, in general, there are not major differences between the data of suspicious and reportedly naïve participants" (p. 804).

But are the effects really "negligible"? At least some observational data suggest this may not be so. Take, for example, the following incident. In the middle of a mock jury study, one of the six jurors experienced a genuine epileptic seizure reminiscent of the feigned seizure that served as a manipulation in a classic study by Darley and Latané (1968). The experimenters, MacCoun and Kerr (1987), reported that "three of the five subjects spontaneously reported that they had questioned the authenticity of the attack" (p. 199), and that "there were indications that prior knowledge of psychological research, derived primarily from course work, was related to suspicion" (p. 199). While the only person who promptly came to the victim's aid had no prior psychology

coursework, “two of the other bystanders reported looking for hidden observers even as they joined in administering aid” (p. 199). The interesting implication of this observation is that had MacCoun and Kerr’s study been concerned with altruistic behavior, then the participants’ behavior, that is, withholding help because they were suspicious of deception and expected to be framed, would have been mistaken as evidence for the “bystander effect” (Darley & Latané, 1968).

Is this just a singular incident in which suspicion compromises experimental data and conclusions? We address this question by looking at three different sets of studies that allow us to examine how suspicion of deception affects behavior in experiments across a wide range of studies. The first set of studies compares conformity behavior of participants who were identified post-experimentally as being either suspicious or unsuspecting of deception. The second set consists of studies that intentionally provoked the expectation of deception at the outset and then examined experimental behavior as a function of it. In the third set of studies, participants’ experimental history (e.g., previous participation in deception studies) was either recorded or systematically manipulated and their experimental behavior studied as a function of it.

To find such studies, we searched for specific keywords in titles and abstracts of articles listed 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 June 2002 (the point at which our analyses were conducted). We also included all studies cited in a recent review by Bonetti (1998a), who concluded that “deception does not appear to ‘jeopardize future experiments’ or ‘contaminate a subject pool’” (p. 389). Finally, we looked up the studies cited in the articles found using the first two methods and included them if they could be classified into one of the three sets.

Set 1: The Effects of Self-Reported Suspicion

To find studies that examined the effects of post-experimentally identified suspicion, we searched for deception in combination with suspicion (and its variants, such as “suspicious”, “suspiciousness”, “suspicions”). This search uncovered two systematic reviews of the social psychology literature that examined the prevalence of suspicion among participants. The studies reviewed by Stricker (1967) excluded with one exception suspicious participants and thus his review does not allow us to examine how suspicion affected experimental behavior.¹⁰ In his review of the literature on social conformity, Stang (1976) found 21 studies that reported the percentage of

“suspicious” participants. Interestingly, for the reported period between 1954 and 1973, the percentage of suspicious participants increased as a function of time, that is, more recent studies observed more suspicious participants ($r = .76$), and about one-third of the studies reported that more than 50% of the participants were suspicious. Thus, one may speculate that the positive correlation between year of study and percentage of suspicious participants tracks the increase of the use of deception in social psychology during that time.

Out of the 21 studies, Stang (1976, p. 363) cited 9 that systematically compared the behavior of suspicious and unsuspecting participants. Typically, this classification was performed on the basis of post-experimental interviews in which participants responded to questions such as “do you feel this experiment was deceptive (involved lying) in any way?” (Geller & Endler 1973, p. 49). In addition to those 9 studies referenced by Stang, our search turned up another 5 studies that examined behavior in experiments as a function of suspicion, all of which were also concerned with conformity behavior. It is probably no coincidence that researchers studying conformity have been particularly concerned with the possible repercussions of suspicion. According to Gross and Fleming (1982), researchers in this area rely heavily on deception, with 96.7% of studies in the area of compliance and conformity having used deception.

As shown in Table 2, in 10 of 14 studies identified by Stang and our additional search, suspicious participants showed less conformity behavior—the target variable in which experimenters were interested—than unsuspecting participants. For 9 of the 10 studies (in which the necessary information was given) we calculated an effect size measure (η^2 , or r); the reduction in conformity as a function of suspicion was of medium to large effect size.¹¹ In four studies (Chipman, 1966; Willis & Willis, 1970; Endler et al., 1972; Wiesenhal et al., 1973) suspicion did not significantly change the amount of conformity behavior, and no study reported that suspicion produced greater conformity.

[Insert Table 2]

To conclude, in research on conformity behavior data of participants who are suspicious of deception and data of naïve participants are different. The observed differences are not “negligible;” in fact they may even represent a lower-bound estimate of the true differences because the classification of suspicious and naïve participants typically rests on people’s self-reports. If one assumes that not all participants reveal their suspicions truthfully, then the resulting misclassification of suspicious participants as naïve participants would have watered down the true differences between

the groups. That participants may not admit their suspicions truthfully is suggested by a number of converging observations. 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. In a similar study, Altemeyer (1971) found that “none of the contaminated Ss reported their foreknowledge or awareness to E” (p. 79; for more recent results see Taylor & Shepperd, 1996).

Set 2: The Effects of Experimentally Induced Suspicion

To circumvent the problem of participants not admitting to being suspicious, experimenters can systematically “plant” participants’ suspicion from the outset and then study their experimental performance as a function of it. To find such studies, we used the search term deception in combination with prebriefing, or forewarning. We found a total of eight studies. The issue with which we are concerned here, namely, the effect of experimentally induced suspicion, was not the explicit focus in all eight studies. Participants’ knowledge and thus suspicion of deception 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 concrete tip-offs by a confederate (e.g., Levy, 1967), to disclosure that deception would occur during the experiment (e.g., Finney, 1987).

Table 3 summarizes how participants’ foreknowledge of deception affected behavior. The results are mixed, with some studies finding no effect and others large effects. Nevertheless, a trend is discernable: When participants received detailed tip-offs about the true purpose of the experiment (e.g., Levy, 1967; Turner & Simons, 1974), were explicitly told that they would be deceived (Finney, 1987), or explicitly acknowledged awareness of experimental manipulation (Golding & Lichtenstein, 1970), suspicion altered experimental performance (albeit not necessarily on all dependent measures). In contrast, when participants were merely informed that some kind of deception might happen (Allen, 1983; Finney, 1987; Wiener & Erker, 1986) or were told the purpose of the study (without indicating the possibility of deception, Gallo et al., 1973) then their performance did not differ from that of control participants not given this information (but see Spinner et al., 1977).

[Insert Table 3]

Why does concrete versus general foreknowledge result in different behavior? One explanation, suggested by Finney (1987), is that with general foreknowledge the effect of suspicion between experimental and control group participants vanishes because it simply does not discriminate experimental participants from those in the control group (without foreknowledge) who may also harbor suspicions (e.g., due to undergraduate psychology training; see, for example, Higbee, 1978).

Set 3: Does Previous Experience of Deception Evoke Suspicion?

Yet another way to explore the effects of suspicion is to study how participants' experimental history affects experimental performance. To find studies that adopted this approach, we used the search term deception in combination with experimental history and found nine studies. Table 4 summarizes a complex picture of findings. In brief, the results suggest that firsthand experience with deception or manipulation affects performance, whereas mere disclosure of the possibility of deception in psychological experiments does not (Cook & Perrin, 1971; Christensen, 1977, Experiments 1 and 2). Second, Silverman et al. (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. In Cook and Perrin's (1971) research, incidental-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 the past and present experimental situation (Brock & Becker, 1966; Cook et al., 1970).

[Insert Table 4]

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).¹² Page and Scheidt were able to replicate the weapons effect in only one out of three of their experiments, and only in a group of participants who had taken part in a deception experiment within the previous month; participants unfamiliar with psychological experimentation did not exhibit the effect. Turner and Simons (1974; see also Simons & Turner, 1976) challenged Page and Scheidt's results, and

Turner, Simons, Berkowitz, and Frodi (1977) even suggested: “Perhaps the failures to replicate the weapons effect occurred because the researchers used subjects who were not naïve about deception or who were very apprehensive about the impression they might create” (p. 369). 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 naïve, the effects of important independent variables may be obscured” (p. 347).

Do Naïve and Suspicious Participants Really Behave the Same?

According to a key argument made in the defense of the use of deception, the differences in the data of naïve and suspicious participants are negligible. We analyzed three sets of empirical studies that systematically explored the effects of suspicion on behavioral data. Though the effects of suspicion are not invariably strong they can be substantial. First, consider the findings in research on conformity in which the effect of suspicion has most extensively been examined. In more than two-thirds of studies, researchers found evidence (of medium to strong effect size) that naïve and suspicious participants exhibited different amounts of conformity behavior (see Table 2). In a second set of studies, we found that the concrete foreknowledge of deception (e.g., being forewarned or prebriefed) can systematically alter experimental performance (see Table 3). Similarly, people’s previous experiences (i.e., having firsthand experienced deception in past experiments) can affect their behavior in experiments (see Table 4). In evaluating this evidence, it is noteworthy that the studies reviewed do not represent opportunistic samples. Rather, we included each study that was identified through a systematic electronic literature search (using the key words listed above).

Our search also identified studies that did not observe suspicion to compromise data. These studies are as important as those that document biasing effects of suspicion because they allow us to identify when experience of deception compromises behavior in experiments. For instance, by virtue of being similar or dissimilar to previous experiments in which participants experienced deception, experimental scenarios appear to differ in the extent to which they elicit suspicion and second-guessing. In addition, dependent variables appear to differ in the extent to which they provide room for biasing effects of suspicion to occur.

What makes participants suspicious? Although all reviewed studies examined suspicion as a function of firsthand experience with deception, there are other candidate sources of suspicion and mistrust. Lipton and Garza (1978) demonstrated that after participating in deception experiments, students in a typical college participant pool talk about the experiment among themselves (despite being told not to do so), thus contaminating the pool for later runs of the experiment. Rubin and Moore (1971) demonstrated that the number of psychology courses could be even more closely related to suspicion than the number of deception experiments in which people recall having taken part. That is, while being taught classic deception experiments, undergraduates may learn to become suspicious. Similarly, 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 courses ($\eta = .43$). Consequently, Higbee recommended that “if psychologists expect the subjects to believe them, perhaps they should get the subjects at the beginning of the semester” (p. 133).

Finally, let us point out that there are at least two possible links between suspicion and experimental performance. Not only can suspicion directly impinge on behavior, it may also interfere with experimental performance by affecting motivations that are thought to be operative in the participants' minds as they approach the research situation. Such motivations include predilections to enact the good subject role, the obedient subject role, the evaluation-apprehensive role, and the negative subject role (for a review of those motivations see Rosenthal & Rosnow, 1991, chapter 6), respectively. Suspicion could amplify some of these motivations while crowding out others. Consistent with this suggestion, Silverman et al. (1970) concluded that experiencing deception makes people more apprehensive of evaluation.

In addition to altering people's motivations, suspecting deception may also render the interpretation of what it means to play a particular subject role more difficult. Evaluation apprehension, for instance, implies that participants are concerned about being observed and judged, and that they will develop hypotheses about how to win positive evaluation and how to avoid negative evaluation (Rosenberg, 1969). But, how can one succeed in making a good impression if one suspects deception—simply by succumbing to the suspected manipulation, by “figuring out” the deception and thwarting it, or by mocking it?

Argument 4: How (In)dispensable Is Deception?

Deception has been defended as an indispensable strategy of last resort for the study of those facets of behavior that are of great social importance, and for which alternative research methods are either unavailable or would produce invalid data (e.g., Bröder, 1998; Kimmel, 1998; Korn, 1998; Weiss, 2001). By this argument, the costs of not conducting such research (e.g., on conformity, obedience, racial stereotypes, bystander effect, and aggression) outweigh the costs of using deception (e.g., Trice, 1986). This argument is particularly compelling because it is explicitly endorsed by the American Psychological Association (APA) rules of conduct. According to those rules

psychologists do not conduct a study involving deception unless they have determined that the use of deceptive techniques is justified by the study's prospective scientific, educational, or applied value and that equally effective alternative procedures that do not use deception are not feasible. (American Psychological Association, 1992, p. 1609)

In other words, the APA rules endorse deception as a strategy of last resort to be used only if its benefits justify its use and if no alternatives are feasible. In what follows, we discuss whether or not the reality of psychology's research practices conforms to this policy prescription. Note that we are not concerned with the more contentious—yet, in light of the APA rules, purely hypothetical—issue of whether or not psychologists (such as experimental economists) could do without the deception. Instead, we focus on the actual issue at hand, namely, is the self-imposed policy prescription effective in guiding research practices. And if not, what can the discipline do to reconcile everyday research practices with the rules of conduct?

Is Deception Used as a Strategy of Last Resort?

Without doubt psychology's use of deception has changed. According to Rosnow and Rosenthal (1997, p. 114), “many of the seminal studies that were conducted then would be impossible today (e.g., Milgram's obedience studies).” That is, the stricter rules of conduct and the establishment of institutional review boards (IRBs) have been effective in preventing the kind of severe cases of deception used in various classic (but also many rather mundane deception studies)

of the 1960s. For a short history of the “ten commandments of the APA” see Rosnow and Rosenthal (1997, chap. 6).

Despite being successful in preventing severe and harmful cases of deception (and thereby protecting both participants and experimenters), the rules appear to have failed in enforcing deception as a strategy of last resort. In fact, mundane deception (not the severe kind of deception that characterize some of the classic deception studies) seems still a routine tool in psychology’s daily research practice. Remember that, for instance, at least every third study published in the prestigious JPSP in the 1990s employed deception. Did all those many studies in this and other social psychology and marketing journals reserve deception for those cases in which the study’s prospective utility justified the use of deception and in which equally effective alternative procedures were not feasible?

We conjecture that in many contemporary deception studies the “no equally effective alternative procedures” requirement was not met. A brief look into past research practices supports this conjecture. Even in research traditions in which deception has been considered to be indispensable, alternative research techniques have often been available. This follows logically from Gross and Fleming’s (1982) review of 1,188 journal articles in leading social psychology journals (between 1959 and 1979). This review analyzed the prevalence of deception in 24 research areas in social psychology (e.g., conformity, altruism, impression formation, attitude change). The authors observed a wide variation in how often deception was used in different areas. Researchers in about half of the areas used deception in less than half of all studies. That is, in areas in which deception has often been advocated as indispensable, alternatives must have always been available—they were simply not used. Even in research areas such as conformity and compliance, which according to Gross and Fleming (1982) used deception in 96.7% of all studies, alternatives are available. This is suggested by Stricker et al.’s (1969) observation that 20% of all conformity studies published in 1964 in four leading social psychology journals did not use deception! Admittedly speculative, one may argue that the use of deception has become the matter of course not because other procedures were not available but because it proved particularly convenient and easy to implement (Hertwig & Ortmann, 2001).

Another way to directly test the conjecture that deception is used when it is dispensable is to examine individual deception studies and see whether or not alternatives were available. We will not reference individual studies because we do not want to single out individual researchers. The issue

here is not the individual behavior but the fact that there is a gap between the ideal stipulated by the APA rules of conduct and the reality of everyday research practice. Consider the following study, recently published in a social psychological journal, which had participants play a dictator game or an ultimatum game. Participants were falsely told that they would be paired with one of the other participants, that on the basis of a chance procedure they alone were assigned the role of the allocator (who had to divide a certain amount of money), and that their income would be contingent on the allocator's and/or recipients' decisions. In addition, the participants were let to believe that the experimenter would hand the allocator's decision to the recipient (involving a rather complicated procedure) and thus decisions would remain anonymous. At the end of the experiment, participants were debriefed, discovering that all participants were allocators and that all received the same amount of money.

This study is not a glaring example of deception. In fact, its lies are rather mundane. Yet, it is an excellent example of the use of deception where it is utterly unnecessary. Experimental economists have conducted countless dictator and ultimatum games without deception. There is nothing in the research question of this particular study that would have required the use of deception. Rather, deception was likely to be motivated by pragmatic consideration such as the need for a larger number of players who are allocators (rather than recipients). Clearly, deception is often less expensive and more convenient than an alternative procedure, but it is only less expensive if it does not entail costs for future experiments: Why would participants who just found out that their decisions and monetary rewards were, contrary to the experimenter's previous claims, not contingent on their behavior trust similar promises in future experiments?

The study described is not a unique case. To assure oneself of this fact, it suffices to browse through recent issues of leading journals in social and experimental psychology. Participants are routinely misled to believe that their decisions in games and gambles will determine their final payoffs, that assignment of roles in an experiment will be determined by chance, that they will be paired up with another person, that the feedback they receive (about their social intelligence, attraction, personality, etc.) will be veridical, and so on. In each of these few, arbitrarily selected examples of the use of deception (taken from contemporary studies), the false claims were not necessary. But why were they then still employed?

How to Narrow the Enforcement Gap

Why are the APA rules of conduct not more effective in enforcing deception as a strategy of last resort? Elsewhere we have argued that one key problem is that the decision of whether deception is justified by its anticipated utility is left to those who stand to benefit from its use (Ortmann & Hertwig, 1997, 1998). Notwithstanding the mediating role of IRBs (which tend to focus on the ethical rather than the methodological consequences of deception), this practice leaves the assessment of private benefits (e.g., relatively quick publication, see Adelson, 1969) and public costs (e.g., contamination of the participant pool) to the interested party (the experimenter)—a classic moral-hazard problem with a solution that currently is not incentive compatible.

How can one enforce the APA rules without necessarily expanding the somewhat daunting role of IRBs, which doubtlessly have complicated the business of experimentation? We believe that the most promising solution to this dilemma is to implement a mechanism in which the individual researcher has an incentive to forgo the routine tool deception and to search and implement alternative procedures. One such pragmatic mechanism has been made possible by recent technological advances: Specifically, we propose that experimenters about to perform deception studies post their experimental designs on an APA website for a specified time period (e.g., two months), thus giving those opposed to deception a chance to suggest workable alternatives. Such an electronic forum would give both defenders and critics of deception the opportunity to discuss on a case-by-case basis whether deception is indispensable (because no equally alternative procedures are feasible). Over time contributions to such a website would generate a depository of successful alternatives and examples of experiments in which they were used so that experimenters considering deception could easily “browse” through alternatives.

How is such a mechanism incentive compatible? To provide an incentive for critics of deception to come up with alternative designs, they should receive credit for it, possibly in the form of co-authorship, in the published articles. In addition, there are also incentives for defenders of deception: By making key aspects of a planned experiment public, these researchers would be “going on record” first with an idea, preempting others following a similar line. In addition, the fact that no alternative procedure was proposed may also help to get fast clearance from the IRBs. That said, the mechanism proposed here is obviously meant to force those who are cavalier about deception to think harder about non-deception alternatives. Given the increased incentive to think harder about alternatives, we expect that few researchers will have to accept co-authors who have bought themselves in through suggestion of a non-deception design. The website solution is thus

meant as an off-equilibrium outcome that rarely happens (see Ortmann & Hertwig, 2002b, for details).

The proposed website is certainly not the only solution to the enforcement gap—others have proposed different measures (e.g., the APA should provide more specific standards regarding the permissibility of deception and its appropriate use; Pittenger, 2002). We hope that the co-occurrence of these proposals (and their underlying concerns) will spur discussion and search for even better mechanisms that have the potential to provide individual researchers with good reasons not to adopt deception “as a matter of course” (Kelman, 1968) but to search for alternative procedures.

Methodological Innovation: A Benefit of Serious Enforcement Efforts?

Enforcing deception as a strategy of last resort would not only promise to minimize the methodological side effects of deception (e.g., suspicion, distrust) but would also provide an incentive to develop, evaluate, and employ alternatives. More than three decades ago, Baumrind (1971) anticipated methodological innovations once deception would no longer be easily available. As she put it:

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. (p. 893)

It is instructive and supportive of Baumrind’s optimism to observe how prohibition of deception in experimental economics appears to spur innovation. Whereas deceiving participants is generally taboo among experimental economists, there are rare dissenting views. One of those dissenters is Bonetti (1998a), who referred to one of the very few deception studies in (experimental) economics to argue that some discoveries require the use of deception. The study in question investigated how much people invest in public goods as a function of the action of other players. Interestingly, this scenario mirrors that of conformity experiments. Specifically, Weinman (1994) had participants divide an endowment (in each round of a repeated game) between a public

and a private good. They made their decisions falsely believing that they interacted either with highly cooperative or very uncooperative players. Weinman observed an asymmetric response to the contribution of others: Players responded to fabricated non-cooperative behavior with non-cooperation but responded to highly cooperative behavior with normal play. According to Bonetti (1998a, p. 387), “the discovery of this evidence supporting what Weinman calls ‘exploitation aversion’ necessarily required deception.”

But does it? Taking up the challenge of developing a procedure that foregoes deception yet yields similar results, Bardsley (2000) developed a design in which he replaced the deceptive scenarios (e.g., high vs. low contribution of others) with fictitious scenarios. The fictitious scenarios were presented side-by-side with the real contributions of other players.¹³ Bardsley showed that under certain conditions participants have incentive to respond to all scenarios as if they were true, thus enabling the experimenter to study scenarios of special interest, such as a high amount of free riding. Confirming Weinman’s (1994) results, Bardsley found that participants appear to be more sensitive to free riding than to cooperation on the part of others.

Clearly, Bardsley’s (2000) alternative to deception is tailored to a specific research question and thus cannot serve as a general alternative template to deception. Yet, this example illustrates that one ought to be careful in reaching categorical conclusions such as that deception is indispensable for the discovery of certain phenomena. Moreover, this case study from experimental economics suggests that abandoning deception as a matter of course, and thus taking the APA guidelines seriously, promises to spur methodological innovation, which in turn will provide researchers with a richer repertoire of alternative non-deception procedures.

General Discussion

In the defense of deception, four arguments are often advanced that we have evaluated on the basis of a systematic search of evidence. Consistent with the first argument, there has been a drop in the use of deception in social psychology (but not in marketing research). Despite this drop, the absolute prevalence in the 1990s was still high in social psychology, and similarly, in marketing research. Second, there is mixed evidence regarding the question of whether deceived participants become resentful about having been fooled. Third, in contradiction to the argument that the data of suspicious and naïve participants do not differ, we found that suspicion has the potential to alter experimental data substantially. Finally, in response to the argument that deception is indispensable

for achieving experimental control in some socially significant areas of research, we argue that deception is frequently used when, in fact, it would have been dispensable.

Undoubtedly, the available evidence is not as clear-cut as either the proponents or the critics of the use of deception sometimes imply. Thus, there is room for honest difference in evaluating the evidence. In what follows, we suggest one way of interpreting the data, and more generally, the manifold costs of deception. While one may certainly disagree with our reading, we believe that it is indisputable that psychology's self-imposed rules of conduct and research reality are two different animals. Unless we mean those rules to be merely cosmetic, the discipline ought to address the gap between them and reality.

Why Bother If Effects of Suspicion Are Negligible:

The Curious Changes in Subject-pool Composition

Possibly, the most important argument in the defense of deception is the conjecture that there are only negligible differences between the data of suspicious and reportedly naïve participants (e.g., Kimmel, 1998, p. 804; Bonetti, 1998a). The data reported in Tables 2-4 contradict this claim. They suggest that firsthand experience with deception can generate suspicion, which, in turn, can alter experimental performance. In contrast, the consequences of secondhand experience are much less clear. To reiterate, secondhand experience with deception is assumed to originate from sources such as undergraduate psychology classes, campus scuttlebutt, media coverage of psychological research, and the profession's reputation more generally. Our search happened upon very few attempts to explore the effects of secondhand experience. In fact, all studies listed in Tables 2-4 examined firsthand experience. Clearly, secondhand experiences are difficult to measure, quantify, and manipulate; they may therefore simply not be studied. Consequently, there is little ground either to corroborate or refute the belief that secondhand experiences contaminate the subject pool.

Once again, are there really no differences between the data of naïve and suspicious participants? A key institutional arrangement in psychology—the drastic shift in subject-pool composition over the last three decades—reflects, in our view, researchers' doubts that suspicion is innocuous. In psychology, undergraduates have been a major source of research data, and, typically, undergraduate participation is enforced through the use of a subject pool. Subject pools are replenished by requiring undergraduate students—notably students from introductory classes—to participate in research projects as part of their course requirements. In a survey of 242 U.S.

psychology departments (with subject pools and graduate programs), Sieber and Sacks (1989) found that 93.4% of departments recruited from introductory courses.¹⁴

Subject pools, however, have not always relied so heavily on students from introductory classes. In his analysis of the participant selection in studies published in the period 1966-1967 in the two largest journals of the APA, Schultz (1969) found that 41% (Journal of Experimental Psychology) and 34% (JPSP) of studies relied on students from introductory psychology courses as participants. What has prompted this drastic change in the composition of psychology's subject pools? Although this is speculation, we suggest that the current practice of recruiting participants mostly from introductory courses could be the result of an evolutionary process—a process driven by attempts to minimize the contaminating effects of deception and suspicion on the participant pool. By replenishing subject pools with ever new and naïve participants and using them as the prime source of data, psychology has designed a recruiting mechanism that promises to curtail the possibly distorting side effects of firsthand experience with deception. If true, researchers may have done nothing less than take Silverman et al.'s (1970) advice to heart, namely, “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).

Psychologists also appear to take individual precautions to curtail the negative consequences of participants' suspicion. For instance, a prominent social psychologist told us that at his laboratory, in which deception is used and in which students are eligible to participate in multiple experiments, experimenters routinely probe for suspicion at the end of the studies. In addition, they ask the participants to list the previous studies they have participated in. If experimenters need naïve participants, they can discount all data from participants who have previously participated in a deception study. Or they might choose to analyze those data separately and estimate the effects of experience with prior deceptions. To the best of our knowledge, this practice is not institutionalized throughout psychological laboratories but is left to the discretion of the individual researcher. Therefore, such arrangements may have the unfortunate, and paradoxical, consequence that researchers who do not use deception are more likely to become victim of its potentially distorting side effects, since they might be less inclined to probe their participants for suspicion and thus be less able to control for the effect of prior experience with deception.

To conclude, in his essay on the potential long-term costs of deception, Kelman (1967) predicted that as we continue to use deception “our potential subjects become increasingly

distrustful of us,” and therefore the “more research we do, the more difficult and questionable it becomes” (p. 7). Was he right? We suggest that firsthand experience with deception indeed has the potential to evoke participants’ suspicion, which in turn can affect data. This potential, however, does not necessarily make future research questionable. Kelman did not anticipate what is possibly an institutional solution to the problem of a contaminated subject pool—psychology’s strategy of constantly replenishing the subject pool, thus reducing the risk of relying on suspicious participants. Economists’ adage that there is no such thing as a free lunch seems to apply to the current context with force. Heavy reliance on a narrow subject pool exacts costs as well. Criticizing social psychology’s heavy dependence on college students as research participants, Sears (1986) argued that “overdependence on this one narrow data base may have unwittingly led us to a portrait of human nature that describes rather accurately the behavior of American college students in an academic context but distorts human social behavior more generally” (p. 515). Judging from Sieber and Sacks’ (1989) results, (social) psychology’s data base appears even narrower—college students from introductory classes—than Sears assumed.

Old but Not Obsolete Evidence and Private Observations

In our search for studies that examined the methodological consequences of deception, we discovered that most available studies date back to the decade between the mid-1960s and the mid-1970s. This is no coincidence. Silverman (1978, p. 405) referred to this period as the “most self-critical decade” of psychology, during which much research was devoted to investigating the “threats to validity that reside in ... the interaction between the experimenter and the subject” (Rosenthal & Rosnow, 1991, p. 110). Are the results of this research obsolete today? For several reasons, we do not think so. For one, although the use of deception has dropped, it is, as we have documented, still frequently used. To the extent that participants’ expectations and degree of suspicion reflect contemporary research practice, the threat of data contamination remains. To the extent that participants’ expectations and degree of suspicion track past research practice, for instance, through undergraduate teaching, the threat of data contamination is even more pronounced.

Second, the few recent studies that are available also indicate the potential of firsthand experience to affect behavior in future experiments. Krupat and Garonzik (1994) and Epley and Huff (1998) asked participants to report what their concrete expectations would be if they participated in future research (e.g., “you will be misled or deceived in some way during the course

of the study”) and analyzed these expectations as a function of prior experience with deception. Participants’ responses suggested that with previous exposure to deception, participants were more likely to expect to be misled and deceived in future experiments, and to be more suspicious of information presented by the experimenter.

Krupat and Garonzik’s (1994) and Epley and Huff’s (1998) findings are also consistent with still another category of contemporary evidence that only rarely makes it to the public domain: researchers’ unprompted, informal observations. One example of such an unprompted observation is MacCoun and Kerr’s (1987) report described earlier. Is theirs just a rare exception or the tip of an iceberg? We do not know. We are, however, surprised by how many of our colleagues have related unprompted observations to us, ranging from comments on participants’ distrust about the promised performance-contingent payment to their distrust of crucial parameters in gambles to their conviction that some coincidental noise outside of the laboratory room is systematically related to the current experiment. These informal observations suggest that there are myriad ways in which suspicion can seep into our labs and studies.

Conclusion

In the late 1960s, Kelman worried that psychologists use deception without question, and he felt that “we are training a generation of students who do not know that there is any other way of doing experiments in our field—who feel that deception is as much *de rigueur* as significance at the .05 level” (1967, p. 3). Since then, some things have changed. Today, we certainly do not teach students that deception is *de rigueur*. Rather, implicitly or explicitly we teach them that deception is a commonly accepted practice that needs, however, to be justified to what is often perceived as a capricious and overly cautious ethical review by IRBs. We doubt that they learn—for instance, by example—that deception is meant to be a strategy of last resort. Moreover, we continue to teach students just one model of investigative situation. In this model, experimenters assume participants’ naiveté, and thus believe they need to camouflage the research purpose—an assumption that appears absurd, for instance, when viewed from the Wundtian model of experimentation (see Danziger, 1990). And finally, we have not begun to teach students to explore alternative approaches actively, nor do we provide them (or us for that matter) with incentives to do so. In this sense, reconsidering the use of deception also affords us the opportunity to both revisit the social

parameters of the investigative situation as well as the implicit and explicit incentives that drive our methodological proclivities.

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., & Lindsay, R. C. L. (1985). Ethical regulations and their impact on research practice. American Psychologist, 40, 59-72.
- Adelson, J. (1969). Personality. Annual Review of Psychology, 20, 136-252.
- Aguinis, H., & Handelsman, M. M. (1997). Ethical issues in the use of the Bogus Pipeline. Journal of Applied Social Psychology, 27, 557-573.
- Aitkenhead, M., & 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.
- Allen, V. L. (1966). Effect of knowledge of deception on conformity. Journal of Social Psychology, 69, 101-106.
- 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.
- Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. Psychological Monographs, 70, (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., & Le Page, A. (1967). Weapons as aggression-eliciting stimuli. Journal of Personality and Social Psychology, 7, 202-207.
- Boles, T. L., Croson, R. T. A. & Murnighan, J. K. (2000). Deception and retribution in repeated ultimatum bargaining. Organizational Behavior and Human Decision Processes, 83: 235-259.
- Bonetti, S. (1998a). Experimental economics and deception. Journal of Economic Psychology, 19, 377-395.
- Bonetti, S. (1998b). Reply to Hey and Starmer & Mc Daniel. Journal of Economic Psychology, 19, 411-414.
- Brandts, J., & Charness, G. (2002). Truth or consequences: An experiment. Manuscript submitted for publication.
- Brock, T. C., & Becker, L. A. (1966). "Debriefing" and susceptibility to subsequent experimental manipulations. Journal of Experimental Social Psychology, 2, 314-323.
- Bröder, A. (1998). Deception can be acceptable. American Psychologist, 53, 805-806.
- Carlson, R. (1971). Where is the person in personality research. Psychological Bulletin, 75, 203-219.
- Chipman, A. (1966). Conformity as a differential function of social pressure and judgment difficulty. 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.
- Clark, R. D. & Word, L. E. (1974). Where is the apathetic bystander? Situational characteristics of the emergency. Journal of Personality and Social Psychology, 29, 279-287.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). [City]: Erlbaum.
- Collins, F. L., Jr., Kuhn, I. F., Jr., & King, G. D. (1979). Variables affecting subjects' ethical ratings of proposed experiments. Psychological Reports, 44, 155-164.

- Cook, T. D., Bean, J. R., Calder, B. J., Frey, R., Krovetz, M. L., & 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. & 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.
- Cooper, J. (1976) Deception and role-playing: On telling the good guys from the bad guys. American Psychologist, 31, 605-610.
- Coulter, X. (1986). Academic value of research participation by undergraduates. American Psychologist, 41, 317.
- Croson, R., Boles, T., and Murnighan, J. K. (in press). Cheap talk in bargaining experiments: lying and threats in ultimatum games. Journal of Economic Behavior and Organization.
- Danziger, K. (1990). Constructing the subject: Historical origins of psychological research. New York: Cambridge University Press.
- Darley, J., & Latané, B. (1968). Bystander intervention in emergencies. Diffusion of responsibility. Journal of Personality and Social Psychology, 8, 377-383.
- Davis, D. D., & Holt, C. A. (1993). Experimental economics. Princeton, NJ: Princeton University Press.
- Endler, N. S., & Hartley, S. (1973). Relative competence, reinforcement and conformity. European Journal of Social Psychology, 3, 63-72.
- Endler, N. S., Wiesensthal, D. L., & 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., & Huff, C. (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., & Silverstein, S. J. (1973). The experimental contract: Subjects' expectations of and reactions to some behaviors of experimenters. American Psychologist, 28, 212-221.
- Erikson, K. (1996). Commentary. American Sociologist, 26, 4-11.

- Ettinger, R. F., Marino, C. J., Endler, N. S., Geller, S. H., & Natziuk, T. (1971). The effects of agreement and correctness on relative competence and conformity. Journal of Personality and Social Psychology, 19, 204-212.
- Farr, J. L., & Seaver, W. B. (1975). Stress and discomfort in psychological research: Subjects' perception of experimental procedures. American Psychologist, 30, 770-773.
- Fillenbaum, S. (1966). Prior deception and subsequent experimental performance: The "faithful" subject. Journal of Personality and Social Psychology, 4, 332-337.
- Fillenbaum, S., & 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., & 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., & Mumford, S. (1973). Effects of deceiving subjects upon experimental results. The Journal of Social Psychology, 89, 99-107.
- Geller, S. H., & 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., & 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. (1979). College students' reactions to social psychological experiments involving deception. Journal of Social Psychology, 107, 99-110.
- Glinski, R. J., Glinski, B. C., & 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., & 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., & 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.
- Herrera, C. D. (1996). An ethical argument against leaving psychologists to their statistical devices. Journal of Psychology, 130, 125-130.
- Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A challenge for psychologists? Behavioral and Brain Sciences, 24, 383-451.
- Hertwig, R., & Ortmann, A. (2002). Economists' and psychologists' experimental practices: How they differ, why they differ and how they could converge. In I. Brocas & J. D. Carillo (Eds.), The psychology of economic decisions (pp. 253-272). New York: Oxford University Press.
- Hey, J. D. (1991). Experiments in economics. Oxford: Basil Blackwell.
- 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.
- Kelman, H. C. (1967). Human use of human subjects: The problem of deception in social psychology. Psychological Bulletin, 67, 1-11.
- Kelman, H. C. (1968). A time to speak: On human values and social research. San Francisco: Jossey-Bass.
- Kimmel, A. J. (1996). Ethical issues in behavioral research: A survey. Cambridge, MA: Blackwell.
- Kimmel, A. J. (1998). In defense of deception. American Psychologist, 53, 803-805.
- Kimmel, A. J. (in press). Ethical trends in marketing and psychological research. Ethics & Behavior.
- Kimmel, A. J., & Smith, N. C. (2001). Deception in marketing research: Ethical, methodological, and disciplinary implications. Psychology and Marketing, 18, 663-689.
- Korn, J. H. (1987). Judgments of acceptability of deception in psychological research. Journal of General Psychology, 114, 205-216.
- Korn, J. H. (1998). The reality of deception. American Psychologist, 53, 805.
- Krupat, E. (1977). A re-assessment of role playing as a technique in social psychology. Personality & Social Psychology Bulletin, 3, 498-504.
- Krupat, E., & Garonzik, R. (1994). Subjects' expectations and the search for alternatives to deception in social psychology. British Journal of Social Psychology, 33, 211-222.

- Lawson, E. (2001). Informational and relational meanings of deception: Implications for deception methods in research. Ethics & Behavior, 11, 115-130.
- Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. Kagel & A. E. Roth (eds.), Handbook of experimental economics (pp. 111-194). Princeton, NJ: Princeton University Press.
- Levenson, H., Gray, M. J., & Ingram, A. (1976). Research methods in personality five years after Carlson's survey. Personality & Social Psychology Bulletin, 2, 158-161.
- Levy, L. (1967). Awareness, learning and the beneficent subject as expert witness. Journal of Personality and Social Psychology, 6, 363-370.
- Lewicki, R. J., & Stark, N. (1996). What is ethically appropriate in negotiations: An empirical examination of bargaining tactics. Social Justice Research, 9, 69-95.
- Lipton, J. P., & Garza, R. T. (1978). Further evidence for subject pool contamination. European Journal of Social Psychology, 8, 535-539.
- MacCoun, R. J., & Kerr, N. L. (1987). Suspicion in the psychological laboratory: Kelman's prophecy revisited. American Psychologist, 42, 199.
- McDaniel, T., & Starmer, C. (1998). Experimental economics and deception: A comment. Journal of Economic Psychology, 19, 403-409.
- McNamara, J. R., & Woods, K. M. (1977). Ethical considerations in psychological research: A comparative review. Behavior Therapy, 8, 703-708.
- Menges, R. J. (1973). Openness and honesty vs. coercion and deception in psychological research. American Psychologist, 28, 1030-1034.
- 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., & 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., & Hertwig, R. (1997). Is deception acceptable? American Psychologist, 52, 746-747.
- Ortmann, A., & Hertwig, R. (1998). The question remains: Is deception acceptable? American Psychologist, 53, 806-807.
- Ortmann, A., & Hertwig, R. (2002a). The empirical costs of deception: Evidence from psychology. Experimental Economics, 5, 111-131.
- Ortmann, A., & Hertwig, R. (2002b). Getting serious about deception: An incentive-compatible mechanism. Manuscript submitted for publication.
- Page, M. M., & 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.
- Pihl, R. O., Zacchia, C., & Zeichner, A. (1981). Follow-up analysis of the use of deception and aversive contingencies in psychological experiments. Psychological Reports, 48, 927-930.
- Pittinger, D. J. (2002). Deception in research: Distinctions and solutions from the perspective of utilitarianism. Ethics & Behavior, 12, 117-142.
- Riecken, H. W. (1962). A program for research on experiments in social psychology. In N. F. Washburne (ed.), Decisions, values, and groups (Vol. 2). New York: Pergamon Press.
- Ring, K. (1967). Experimental social psychology: Some sober questions about some frivolous values. Journal of Experimental Social Psychology, 3, 113-123.
- Ring, K., Wallston, K., & Corey, M. (1979). Mode of debriefing as a factor affecting subjective reaction to a Milgram-type obedience experiment: An ethical inquiry. Representative Research in Social Psychology, 1, 67-88.
- Rosenberg, M. J. (1969). The conditions and consequences of evaluation apprehension. In R. Rosenthal & R. L. Rosnow (Eds.), Artifact in behavioral research (pp. 279-349). New York: Academic Press.
- Rosenthal, R., & Rosnow, R. L. (1991). Essentials of behavioral research: Methods and data analysis (2nd ed.). Boston: McGraw Hill.
- Rosnow, R. L., & Rosenthal, R. (1997). People studying people: Artifacts and ethics in behavioral research. New York: Freeman.
- Rubin, Z., & Moore, J. C., Jr. (1971). Assessment of subjects' suspicions. Journal of Personality and Social Psychology, 17, 163-170.

- Schultz, D. P. (1969). The human subject in psychological research. Psychological Bulletin, 72, 214-228.
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on psychology's view of human nature. Journal of Personality and Social Psychology, 51, 515-530.
- Seeman, J. (1969). Deception in psychological research. American Psychologist, 24, 1025-1028.
- Sharpe, D., Adair, J. G., & 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., & Rodriguez, B. (1995). Deception methods in psychology: Have they changed in 23 years? Ethics and Behavior, 5, 67-85.
- Sieber, J. E., & Saks, M. J. (1989). A census of subject pool characteristics and policies. American Psychologist, 44, 1053-1061.
- Silverman, I., Shulman, A. D., & 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, C. P. (1981). How (un)acceptable is research involving deception? IRB: A review of Human Subjects Research, 3, 1-4.
- Smith, C. P., & Berard, S. P. (1982). Why are human subjects less concerned about ethically problematic research than human subject committees? Journal of Applied Social Psychology, 12, 209-222.
- Smith, S. S., & 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.
- Spinner, B., Adair, J. G., & 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., & 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., & Jackson, D. N. (1967). Suspicion of deception: Implication for conformity research. Journal of Personality and Social Psychology, 5, 379-389.
- Stricker, L. J., Messick, S., & Jackson, D. N. (1969). Evaluating deception in psychological research. Psychological Bulletin, 71, 343-351.
- Sullivan, D. S., & Deiker, T. A. (1973). Subject-experimenter perceptions of ethical issues in human research. American Psychologist, 28, 587-591.
- Taylor, K. M., & Shepperd, J. A. (1996). Probing suspicion among participants in deception research. American Psychologist, 51, 886-887.
- Tesch, F. E. (1977). Debriefing research participants: Though this be method there is madness to it. Journal of Personality and Social Psychology, 35, 217-224.
- Toy, D., Olsen, J., & Wright, L. (1989). Effects of debriefing in marketing research involving "mild" deceptions. Psychology & Marketing, 6, 69-85.
- Trice, A. D. (1986). Ethical variables? American Psychologist, 41, 482-483.
- Turner, C. W., & 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., & 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.
- Weber, S. J., & Cook, T. D. (1972). Subjects effects in laboratory research: An examination of subject roles, demand characteristics, and valid inferences. Psychological Bulletin, 77, 273-295.
- Weinman, J. (1994). Individual behavior in a free-riding experiment. Journal of Public Economics, 54, 185-200.
- Weiss, D. J. (2001). Deception by researchers is necessary and not necessarily evil. Behavioral and Brain Sciences, 24, 431-432.
- Wiener, R. L., & 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., & 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., & Willis, Y. (1970). Role playing versus deception: An experimental comparison. Journal of Personality and Social Psychology, 16, 472-477.

Authors' Note

Ralph Hertwig, Max Planck Institute for Human Development, Center for Adaptive Behavior and Cognition, Berlin, Germany, and Andreas Ortmann, CERGE-EI, a joint workplace of Charles University and the Academy of Sciences of the Czech Republic, Prague. We would like to thank Valerie M. Chase, Gerd Gigerenzer, Anita Todd, Peter M. Todd, Tom Wallsten, and David Weiss for many constructive comments. Correspondence should be addressed to Ralph Hertwig, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. Electronic mail may be sent to hertwig@mpib-berlin.mpg.de.

Table 1. A sample of conclusions from psychologists and experimental economists regarding the negative effects of deception.

Orne (1962, pp. 778-779). [The use of deception] on the part of psychologists is so widely known in the college population that even if a psychologist is honest with the subject, more often than not he will be distrusted. As one subject pithily put it, "Psychologists always lie!" This bit of paranoia has some support in reality.

Ring (1967, p. 118). What is the perceptive student to think, finally, of a field where the most renowned researchers apparently get their kicks from practicing sometimes unnecessary and frequently crass deceptions on their unsuspecting subjects?

The short-run gains may be considerable, but it does not appear chimerical to suggest that the ultimate price of deception experiments may be the creation of extremely mistrustful and hostile subject pools. It would be ironic indeed if, by their very style of research, social psychologists were to put themselves out of business.

Argyris (1968, p. 187). Many experiments have been reported where it was crucial to deceive the students. . . . One result that has occurred is that students now come to experiments expecting to be tricked. The initial romance and challenge of being subjects has left them and they are now beginning to behave like lower level employees in companies. Their big challenge is to guess the deception (beat the management). If one likes the experimenter, then he cooperates. If he does not, he may enjoy botching the works with such great skill that the experimenter is not aware of this behavior.

Adelson (1969, p. 220). When the campus population learns, as it can hardly fail to do, about the common tendency of psychologists to deceive, so that all kinds of unanticipated, unknown expectations enter the experimental situation, the subject aiming to "p sych" the experimenter's "psyching" of him, subject and experimenter entangled in a web of mutual suspicion, mutual deception.

Seeman (1969, pp. 1025-1026). When a subject has once participated in a study using deception he is no longer a naive subject but a sophisticated subject who brings to subsequent studies a variety of personal theories and hypotheses that guide the behavior of the subject quite as decisively as theories and hypotheses guide the behavior of an experimenter. In view of the frequency with which deception is used in research we may soon be reaching a point where we no longer have naive subjects, but only naive experimenters. It is an ironic fact that the use of deception, which is intended to control the experimental environment, may serve only to contaminate it.

Hey (1991, pp. 21, 119, 173, 225). I feel that it is crucially important that economics experiments actually do what they say they do and that subjects believe this. I would not like to see experiments in economics degenerate to the state witnessed in some areas of experimental psychology where it is common knowledge that the experimenter says one thing and does another. [Subjects] believing what the experimenter tells them...seems to me to be of paramount importance: once subjects start to distrust the experimenter, then the tight control that is needed is lost.

Davis and Holt (1993, pp. 23-24). The researcher should...be careful to avoid deceiving participants. 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. 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.

Ledyard (1995, p. 134). It is believed by many undergraduates that psychologists are intentionally deceptive in most experiments. If undergraduates believe the same about economics, we have lost control. It is for this reason that modern experimental economists have been carefully nurturing a reputation for absolute honesty in all their experiments.

Table 2. The effects of suspicion on experimental performance.

Authors	Proportion of suspicious participants	Experimental performance of suspicious participants (effect size ^a)
Allen, 1966	30 out of 120 (25%)	Less conformity: On a maximum score of 100% conformity, unsuspecting participants scored on average 26% and suspicious participants 12%
Stricker, Messick, & Jackson, 1967	38.6% (averaged across sex and suspicion about various aspects of the experiment; see their Table 1)	Less conformity: $r = .49$ (Table 4), $r = .33$ (Table 5); averaged across sex and measures of conformity
Glinski, Glinski, & Slatin, 1970	Sessions 1 and 2: 42 out of 55 (76%)	Less conformity: $r = .89$ (Session 1), $r = .86$ (Session 2)
Ettinger, Marino, Endler, Geller & Natziuk, 1971	15 out of 40 (38%)	Less conformity: $\eta = .33$
Endler, Wiesenthal & Geller, 1972	No data	No difference in conformity (no “significant” main effect)
Endler & Hartley, 1973	14 out of 40 (35%)	Less conformity: $\eta = .31$
Geller & Endler, 1973	28 out of 54 (52%)	“Once subjects become suspicious,

Deception and Experimental Control 43

		their conformity sharply decreases” (p. 52): $\eta = .6$
Geller, Endler & Wiesenhal, 1973	21 out of 61 (34%)	Less conformity: $\eta = .33$
Wiesenhal, Endler & Geller, 1973	96 out of 116 (83%)	No difference in conformity (non-significant t-test)
Chipman, 1966 ^b	19 out of 68 (28%)	No significant difference in conformity
Willis & Willis, 1970 ^b	54.2%	Little to no effect
Rubin & Moore, 1971 ^b	95 out of 142 (67%) were either medium or highly suspicious	Less conformity: $r = -.42$
Adair, 1972 ^b	38 out of 86 (44%)	Less conformity: $\eta = .21$
Stang, 1976 ^b	13 of 65 (20%)	Less conformity: $\eta = .3$; “significant’ treatment effects on conformity only when suspicious Ss were removed from the analyses” (p. 353)

^a Effect sizes calculated (η , biserial correlation r) when sufficient information was available.

^b Obtained from our literature search (search words: “deception” and “suspicion” and its variants); articles with no index stem from Stang’s (1976) review.

Table 3. The effects of anticipation of deception on experimental performance.

Authors	Research topic	Manipulation	Behavioral effects (Effect size ^a)
Levy, 1967	Verbal conditioning	Two groups of participants: fully informed (tipped-off by a confederate) and uninformed	Groups differed in the level of performance ($\eta = .41$) but there were no significant differences in the shape of the acquisition curve
Golding & Lichtenstein, 1970	Valins effect (effect of bogus heart rate feedback on preferences)	Three groups of participants: naïve, suspicious (by being told in a conversation with a confederate that that they would be tricked), and completely informed about the deception by a confederate	No “significant” differences in the Valin effect as a function of prior knowledge. However, participants who admitted awareness of experimental manipulation in a postexperimental questionnaire did not show the Valins effect, whereas those who either were not aware of or did not admit their awareness showed a substantial effect ($r = -.48$)
Gallo, Smith, & Mumford, 1973	Conformity behavior	Three groups of participants: complete, partial, or no information about the purpose of the experiment (the information did not reveal that deception was used)	No significant effect ($\eta = .13$)

Deception and Experimental Control 45

<p>Turner & Simons, 1974</p>	<p>Aggression (weapons effect)</p>	<p>Three groups of participants: no information, informed that some deception might be involved (by a confederate tip-off), or informed that “the weapons were probably part of the procedure to influence their behavior” (p. 342)</p>	<p>“Increased levels of ... subject sophistication led to decreased numbers of shocks administered by subjects to their frustrators” (p. 341; <u>eta</u> = .43)</p>
<p>Spinner, Adair, & Barnes, 1977</p>	<p>Incidental learning</p>	<p>At the end of the first part of an experiment, designed to arouse suspicion, participants were told: “Sometimes experiments require that a subject be deceived initially” (p. 546). Based on an “awareness questionnaire” administered at the end of the second part of the experiment, participants were classified into three groups as a function of their suspicion and anticipation of other tasks</p>	<p>Those who were suspicious and intended to prepare for some other task scored higher than those who did not prepare and/or were not suspicious (<u>eta</u> = .46)</p>
<p>Allen, 1983</p>	<p>Cooperativeness in a Prisoner’s Dilemma game</p>	<p>Two groups of participants: “neutral forewarning” (i.e., “in a few experiments it is necessary for experimenters to deceive subjects concerning some elements of the experiment”) vs no forewarning</p>	<p>No significant effect</p>

Finney, 1987	Conformity behavior	Three groups of participants were instructed that they either “may be deceived,” or “will be deceived,” or they did not receive any consent information	The number of conformity judgments in the “will be deceived group” (4.1) was significantly higher than in the “no consent” group (1.9), but the results in the latter group did not differ from those in the “may be deceived group” (2.3)
Wiener & Erker, 1986	Attribution of responsibility and evaluation of culpability	Two groups: standard informed consent, and prebriefing group (i.e., participants were alerted to the possibility that they might be intentionally misinformed)	No significant effects for sentencing judgment, verdicts, and attribution judgments; significant differences in the attribution process

^a Effect sizes calculated when sufficient information was available (search words: “deception” and “prebriefing,” “forewarning,” or “informed consent”).

Table 4. The effects of experimental history on participants' performance.

Authors	Research topic	Manipulation	Behavioral effects (Effect size ^a)
Brock & Becker, 1966	Compliance behavior	Students participated in two consecutive experiments, the debriefed experiment and the test experiment. Participants were assigned to three groups: no debriefing, partial debriefing, and complete debriefing. For half of the participants, the test experiment included an element from the debriefing experience; for the other half the common element was omitted.	Complete debriefing reduced compliance behavior in the test experiment (10%) but only when the debriefing situation and the test experiment were explicitly similar; no reduction in the no and partial debriefing conditions (50% and 50%)
Fillenbaum, 1966	Incidental learning	<u>Experiment 1.</u> Performance on an incidental-learning task after an earlier task that did or did not involve deception. <u>Experiment 2.</u> Same procedure as in	<u>Experiment 1.</u> Although participants who experienced deception did somewhat better on the incidental-learning task, the difference was “not very large and far from significant” (p. 534, $r = .1$). Difference was larger if one compared participants who reported themselves to be suspicious to those who did not.

		<p>Experiment 1 with minor changes.</p>	<p><u>Experiment 2</u>. Participants with deception experience did better on the incidental-learning task ($r = .27$). As in Experiment 1, difference was larger if one compared participants who reported themselves to be suspicious to those who did not.</p>
<p>Cook, Bean, Calder, Frey, Krovetz, & Reisman, 1970</p>	<p>Attitude-change experiments</p>	<p><u>Experiment 1</u>. Experimentally naïve participants took part in one of five attitude-change experiments.</p> <p><u>Experiment 2</u>. Participants were assigned to one of three groups in Experiment 1, which was or was not linked to Experiment 2 (by a common cue). The three groups were no deception, experience of deception,</p>	<p><u>Experiment 1</u>. Attitude data did not significantly differ as a function of experimental history. Experimental history, however, affected global attitudes: Participants with deception experiences believed the experimenter less, considered experiments to be less scientific and less valuable, and reported caring less about understanding and following instructions</p> <p><u>Experiment 2</u>. Attitude was affected by the deception variable and the presence of the cues ($\eta = .34$). Without a cue, experience of deception biased the data</p>

		knowledge of deception.	(compared to knowledge of deception). With a cue, learning about deception but not experiencing it biased the data ^b
Fillenbaum & Frey 1970	Incidental learning	Students were given the critical incidental-learning task immediately after a prior and revealed deception on another task. Students were categorized as “trustful” or “suspicious” participants.	Suspicious participants scored higher on the incidental-learning task than trustful participants (<u>eta</u> = .31)
Silverman, Shulman, & Wiesenthal, 1970	Various dependent variables	Experiment 1 involved either deception and debriefing or a memory study without deception. In Experiment 2, all participants were given tests measuring compliance of demands, persuasibility, sentence completion, and a personality test.	“Significant differences between deception and nondeception conditions were observed with all four of the tests used” (p. 209). <u>Eta</u> equaled .25, .26, and .29 for the compliance of demands, persuasibility, and sentence completion test, respectively. Overall, “the deception experience sensitized subjects to possible ulterior purposes of experiments, increasing evaluation apprehension” (p. 209)

Cook & Perrin, 1971	Attitude change, incidental learning	<p><u>Experiment 1.</u> Participants were assigned to one of three deception conditions: no deception, experience of deception, and knowledge of deception.</p> <p><u>Experiment 2.</u> Attitude-change and incidental-learning measures were obtained (participants did or did not learn that also this experiment involved subsequent deception; we ignore this manipulation here).</p>	<p>The attitude data (unlike in Cook et al, 1970) did not discriminate between conditions. The incidental-learning measure showed that prior experience but not prior learning of deception produced greater incidental learning ($r = .3$), and “experiencing deception produced the strongest evidence of absolute bias” (p. 215). A measure of general suspiciousness (“how truthful are psychology experimenters”) but not of particular suspiciousness (concerning the relationship of both experiments) showed a main effect on incidental learning ($\eta^2 = .29$)</p>
Page & Scheidt, 1971	Aggressiveness (weapons effect)	<p><u>Experiment 3.</u> Two groups of participants: naïve participants who took part in a psychological experiment for the first time and sophisticated participants who took part in a deception experiment within the last month.</p>	<p>The weapon effect was obtained for the sophisticated but not for the naïve participants ($\eta^2 = .32$). “What appeared to be aggressive behavior to the original investigators seems to have been a sham or an artifact” (p. 315).</p>

<p>Christensen, 1977</p>	<p>Verbal conditioning</p>	<p><u>Experiment 1.</u> Four experimental groups, including one group in which an active attempt was made to manipulate their behavior. Then they were debriefed and went through the verbal conditioning procedure.</p> <p><u>Experiment 2.</u> Three experimental groups, including one prior manipulation group and one non-manipulation group</p>	<p><u>Experiment 1.</u> Conditioning did not occur for the group that experienced prior manipulation and deception, but it did occur for the group that was only told that experiments may involve active manipulation of their behavior</p> <p><u>Experiment 2.</u> Unlike in the control and non-manipulation groups, “subjects given a manipulative experimental experience do not exhibit verbal conditioning” (p. 397)</p>
<p>Gruder, Stumpfhauser, & Wyer, 1977</p>	<p>Performance on an intelligence test</p>	<p>Participants received randomly determined feedback about their performance on an intelligence test. Half of them were debriefed about this deception whereas the other half were not. Then they worked on a parallel form of the test a week later.</p>	<p>Participants who had been debriefed improved more in the parallel form than those who had not been debriefed ($\eta^2 = .3$)</p>

^a Effect sizes calculated when sufficient information was available. ^b Bias being defined as the difference to the no deception group.

Footnotes

¹ We note that ethical questions are not divorced from methodological ones. For example, the moral thing to do with deceived participants is to debrief them as quickly as possible after completing the experiment. Each time this moral imperative (which the APA has endorsed) is met, however, the impression that psychologists commonly deceive may be strengthened (Tesch, 1977).

² They are Stricker (1967), Seeman (1969), Carlson (1971), Menges (1973), Levenson, Gray, and Ingram (1976), Krupat (1977), McNamara and Woods (1977), Gross and Flemming (1982), Adair, Duschenko, and Lindsay (1985), Vitelli (1988), Toy, Olsen, and Wright (1989), Sieber, Iannuzzo, and Rodriguez (1995), Nicks, Korn, and Mainieri (1997), and Eply and Huff (1998). Adair et al. (1985), for instance, defined deception as “the provision of information that actively misled subjects regarding some aspect of the study” (p. 62). Nicks et al. (1997) defined deception as an “explicit misstatement of fact” (p. 70), and Menges (1973) described deception as instances where “the subject is given misleading or erroneous information” (p. 1032).

³ There is a completely different approach to the definition of deception. While deception is commonly defined on the basis of the experimenter’s behavior (e.g., intentionally providing false information), one could define it alternatively on the basis of how participants perceive the experimenter’s behavior. According to such a definition, deception would have occurred if participants, after being completely debriefed, had perceived themselves as being misled. Such an approach defines deception empirically and post hoc rather than on the basis of abstract principles. We do not know of any attempt to realize such an “inductive” approach.

⁴ To test the often-heard belief that deception is almost exclusively used in social psychology studies, we conducted a small-scale survey in which we asked researchers in the field of behavioral decision making how often they used deception in their research practice (Hertwig & Ortmann, 2001). A total of 26 researchers responded; the sample encompassed well-established as well as young researchers. They estimated that, on average, they used deception in 1.7 out of 10 experiments. It was particularly interesting to observe that the researchers’ opinions about the use of deception varied widely within one research area. Whereas 60% percent of the researchers stated that they never use it, 20% use it in half or more of their experiments.

⁵ Curiously, Christensen (1988) in his review of the evidence did not include his own study conducted a decade earlier. There, he concluded: “The primary conclusion that can be drawn from the present two studies is that subjects who have knowingly participated in a manipulative experiment will attempt to resist such a manipulative intent in future manipulative experiments” (pp. 399-400).

⁶ It is not without irony that Milgram, who so impressively demonstrated that authority is seductive, defended his experiments in part by data from follow-up questionnaires revealing that the majority of participants were glad they had participated.

⁷ That participants appear to feel less bored in psychological experiments that suffer from a “mundane, repetitive nature” (Sharpe, Adair & Roese, 1992, p. 589; see also Coulter, 1986) is a two-edged benefit: Whereas deception (or just the suspicion thereof) may in fact make a tedious and boring study less so by triggering curiosity and the desire to see through the deception (in the present and even more so in future experiments), these very responses can undermine experimental control.

⁸ To avoid opportunistic sampling of evidence (see Footnote 5), we referenced here all published journal articles that Kimmel (1998, pp. 104-107) in his recent review cited in support of his conjecture that the negative effects of deception appear to be minimal. We also included other articles in his support that we encountered outside of his review. Finally, we attempted to unearth further studies that gauged students’ feeling of resentment (or lack thereof) about the use of deception in psychology experiments. Using the term deception in combination with either feelings, or resentment, we conducted a literature search using PsycINFO/PsycLIT. These searches did not turn up any further hits.

⁹ Incidentally, these studies also demonstrate that an experimenter does not need to deceive in order to experimentally study deception and its consequences.

¹⁰ Stricker (1967) surveyed all articles using deception that were published in four leading social psychology journals in 1964. Of the 88 deception studies, only 16 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) later suggested that “it seems very likely that the overall rate of actual suspicion in these 16 studies was seriously underestimated” (p. 345) due to the inadequacy of the suspicion criteria used. Using different criteria, later studies found much larger numbers of suspicious participants (see our Table 2).

¹¹ The effect size measure eta is defined as the square root of the proportion of variance accounted for (Rosenthal & Rosnow, 1991) and is identical to the Pearson product-moment correlation coefficient when df = 1, as is the case when two conditions are compared (as in most cases where we calculated eta). According to Cohen’s (1988)

classification of effect sizes, a value of η^2 of .1, .3, or .5 constitutes a small, medium, or large effect size, respectively.

¹² The “weapons effect” (originally reported by Berkowitz & Le Page, 1967) suggests that weapons might stimulate aggression by classical conditioning processes resulting from learned associations between aggressive acts and weapons.

¹³ The instruction participants received read as follows: “In each situation, your screen will show you data representing other people’s choices. This data will either be randomly generated, or in some cases, set by the experimenter, unless the situation is the real situation. If so, the numbers will show the actual decisions made by the rest of the group. [...] We would like you to treat each situation as if it is real and the only situation. Note that for all you know, each one could be the real one, in which case ALL information you are given about it is true” (Bardsley, 2000, p. 235).

¹⁴ 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).