Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Cognition xx (xxxx) 1–19
COGNITION
www.elsevier.com/locate/COGNIT

Identification, situational constraint, and social cognition: Studies in the attribution of moral responsibility

Robert L. Woolfolk[a,c,*], John M. Doris[b], John M. Darley[c]

[a]*Department of Psychology, Rutgers University, Piscataway, NJ 08854, USA*
[b]*Philosophy-Neuroscience-Psychology Program, Washington University, St. Louis, MO 61330, USA*
[c]*Department of Psychology, Princeton University, Green Hall, Princeton, NJ 08544-1010, USA*

Received 4 April 2005; revised 25 May 2005; accepted 25 May 2005

Abstract

In three experiments we studied lay observers' attributions of responsibility for an antisocial act (homicide). We systematically varied both the degree to which the action was coerced by external circumstances and the degree to which the actor endorsed and accepted ownership of the act, a psychological state that philosophers have termed 'identification'. Our findings with respect to identification were highly consistent. The more an actor was identified with an action, the more likely observers were to assign responsibility to the actor, even when the action was performed under constraints so powerful that no other behavioral option was available. Our findings indicate that social cognition involving assignment of responsibility for an action is a more complex process than previous research has indicated. It would appear that laypersons' judgments of moral responsibility may, in some circumstances, accord with philosophical views in which freedom and determinism are regarded to be compatible.
© 2005 Published by Elsevier B.V.

Keywords: Social cognition; Responsibility; Attribution; Identification

* Corresponding author.
E-mail address: woolfolk@princeton.edu (R.L. Woolfolk).

0022-2860/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.cognition.2005.05.002

UNCORRECTED PROOF

1. Introduction

The ascription of moral responsibility is ubiquitous in both everyday social interaction and institutionalized social practices. The ways in which people understand and assign responsibility have been of great interest to psychologists and cognitive scientists studying social cognition and the attribution of responsibility, as well as to philosophers working in ethical theory. However, the folk theories that social perceivers employ in ascribing responsibility remain incompletely understood.

Empirical research on folk theories addressing the assignment of responsibility has its basis in attribution theory. Jones & Davis (1965) developed their theory of "correspondent inference" to articulate the conditions under which the observer of another person's actions would believe that those actions "corresponded" with or were indicative of the underlying actor's intentions, attitudes, or traits. They suggested that observers make correspondent inferences only after concluding that an actor is free to choose to perform the observed act, versus being constrained to do so by external factors. Kelley's (1972) "discounting principle" expresses the same conclusion. It maintains that attributions regarding characteristics of an actor, based on observations of that actor's behavior, are made only when the observed behavior is thought to be unconstrained. Conversely, in cases where an act is performed under extreme constraint, inferences about characteristics of the actor are expected to be "discounted."

The empirical research, however, has shown that observers sometimes fail to discount the informational value of behavior that is compelled or coerced. In a long run of "no choice" experiments associated with Jones (1990), participants made correspondent inferences (attributed characteristics to the actor) even when it appeared to be obvious that the actions they observed were produced by strong and visible external constraints. In subsequent research, this tendency toward "overattribution" (Quattrone, 1982) has been demonstrated for a wide variety of attitudes and traits, leading Ross (1977) to coin the now famous "fundamental attribution error" term to describe this effect (cf. Darley & Cooper, 1998; Gilbert & Malone, 1995; Jones, 1990; Ross & Nisbett, 1991).

Recent work (e.g., Malle, 1999; McClure, 1998) has complicated the picture somewhat and has suggested that the person/situation dichotomy of causes, upon which much attribution research is predicated, is an overly simplistic framing of ordinary persons' thinking, and recent empirical studies indicate that overattribution may be less pervasive than suggested by early demonstrations of the "fundamental attribution error" (Fein, Hilton, & Miller, 1990; Hilton, Fein, & Miller, 1993; see Gilbert & Malone, 1995, for discussion). Although this recent research suggests important qualifications regarding the nature and extent of overattribution, the cumulative weight of evidence indicates that when behavior is constrained, perceivers regularly attribute more influence to characteristics of the person, as opposed to properties of the situation, than the discounting principle would predict.

According to various psychological theories (Darley & Shultz, 1990; Shaver, 1985; Shultz & Schleifer, 1983), the personal characteristic of actors that most influences observers' attributions is the actor's perceived *causal role* in an outcome. Some of the more recent psychological accounts of responsibility attribution emphasize the extent to which a given outcome is in the actor's "control" and is intentionally brought about

UNCORRECTED PROOF

(Alicke, 2000; Schlenker, Britt, Pennington, Murphy, & Doherty, 1994; Weiner, 1995). According to Schlenker et al.'s (1994) "triangle model" of responsibility, actors are likely to be held responsible when there is a perceived link between the event and the actor, such that the actor is viewed as having foreseen and "freely" brought about the event. Alicke's (2000) "culpable control" model of blame assignment posits various forms of personal control that are attended to by observers in attributing moral responsibility. One of these is "volitional outcome control,", i.e. the extent to which the observer desired and foresaw the outcome of her action. Similarly, Reeder, Kumar, Hesson-McInnis, & Trafimow (2002) found that participants judging the morality of an actor paid more attention to the actors' motives, such as self-interest, than to "facilitating situational forces." Weiner's (1995) theory of responsibility also emphasizes the attribution of psychological states to the actor in the assessment of credit and blame; Weiner contends that in some circumstances observers may assign responsibility before evaluating mitigating contextual factors. While recent theories retain the familiar emphasis on the actor's causal role, especially causally efficacious psychological antecedents of behavior, these theories recognize that individuals can be connected to actions in highly complex ways that are not well summarized by relatively simplistic attributional principles, such as the discounting principle.

Philosophers often have maintained that individuals should not be held accountable for acting or failing to act when insufficient capacity for appropriate behavior is present, or when operating under constraints they could not be reasonably expected to resist (Smith, 1961; Wallace, 1994). These philosophical intuitions also are evident in legal practice (Robinson, 1996), in that we do not hold children accountable for acts that would constitute criminal conduct in an adult, nor do we court martial military personnel who denounce their country while being tortured as prisoners of war. Indeed, many philosophers have endorsed a principle similar to the psychologist's causal discounting principle: if a behavior is determined by factors outside of the actor's control, the actor is not morally responsible for that behavior.

Here the philosophical discussion of "freedom and determinism" becomes relevant and we will briefly review it. The problem is among the most controversial and recalcitrant in philosophy, and numerous commentators have observed that existing theoretical accounts of responsibility have difficulties, perhaps insoluble difficulties (e.g. Kane, 2002; Nagel, 1986). How can people act freely, it is asked, if, as the "scientific world view" holds, all behavior is causally determined by antecedent forces, forces beyond the actor's control?

There are three standard responses to this question. *Hard determinists* deny that people are ever responsible for their behavior, while *libertarians* insist that causal determinism is not always true in the case of human behavior, thereby allowing for the possibility of moral responsibility. These two groups are labeled *incompatiblists*, in that they both regard causal determination of behavior as incompatible with moral responsibility. Incompatibilists subscribe to the famous "principle of alternate possibilities," which states that one is morally responsible for what one has done only if one could have done otherwise. *Compatibilists*, on the other hand, assert that moral responsibility and causal determinism can be simultaneously maintained and that people may be legitimately held responsible in violation of this principle, even when they could *not* have done otherwise.

UNCORRECTED PROOF

An influential compatibilist approach is associated with the philosopher Harry Frankfurt (1988; cf. Bratman, 1996; Velleman, 1992). Frankfurt contends that judgments of responsibility for behavior should be governed by the extent to which the actor "identifies" with the behavior and the motivations that produce the behavior (Doris, 2002; Frankfurt, 1988). An actor identifies with a behavior (or its motives) when she "embraces" that behavior (or its motives) or performs the behavior "wholeheartedly" (Bratman, 1996; Frankfurt, 1988); we might say that an actor identifies with a behavior to the extent that it expresses her "fundamental evaluative orientation" (Watson, 1996). When Dan happily donates some money to the office charity because he is deeply committed to giving, he identifies with his behavior, and is, therefore, to be credited even if his boss has pressured him to donate. The converse of identification occurs when the actor is "alienated" from the desires or motives associated with the behavior, where the desires seem to result from factors external to the self. For example, when Julie grudgingly and unwillingly contributes to the office charity solely to appease her boss, she does not identify with the act of giving, and therefore is not to be praised for the "charitable" deed.

While the kind of compatibilism described above has considerable intuitive appeal in these cases, it has implications that appear to be strikingly counterintuitive when generalized: specifically, the theory appears to imply that persons may be held responsible even when they operated in the grip of forces compelling that action. In the studies we report in this article, we investigate whether participants attribute elevated responsibility to actors who are coerced to perform actions they also strongly desire to perform.

In the present study we examine empirically the following questions: (1) What is the effect of an actor's degree of identification with an act on observers' attributions of responsibility? (2) What is the effect of situational constraint on observers' attributions of responsibility? (3) To what degree do constraint and actor identification interact as influences on observers' attributions of responsibility? More specifically, does a high desire to commit the act increase attributions of responsibility even when the act seems compelled, or at least strongly coerced by external circumstances?

We use the term "identification" to denote the degree to which an actor wants or desires to perform behavior and maintains a positive "fundamental evaluative orientation" (Watson, 1996) toward that behavior. We investigate the variable of identification together with a more familiar causal factor, the degree to which the actor was coerced or compelled to perform the action. Our hypothesis is that an agent's identification with a behavior influences responsibility attributions to that actor, even when the actor is strongly constrained to do the action. The significance of this hypothesis, if supported, is two-fold. First, it indicates that habits of responsibility attribution are influenced by factors other than the causal/explanatory factors implicated in the theories reviewed above. Second, it suggests that one philosophical account of responsibility, the "identificationist" account associated with Frankfurt, is reflected in the way ordinary people think.

Three experiments are described. In the first two experiments, we systematically varied identification, here operationalized as the extent to which an actor both endorses an action and desires to perform it. Identification was crossed with the level of constraint, or the degree to which the act was coerced. The third experiment includes a check on the independent variable manipulation of constraint employed in the first two experiments.

UNCORRECTED PROOF

2. Experiment 1

In Experiment 1, we assessed the impact of varying levels of identification and situational constraint on the attribution of responsibility for a violent action, the killing of another person. We hypothesized that both external constraint and the actor's level of identification would affect judgments of responsibility for the action, even when the action was highly constrained.

2.1. Method

2.1.1. Participants and procedure

Seventy-two (34 female and 38 male) University of California, Santa Cruz undergraduates enrolled in philosophy classes participated in the experiment as volunteers. Participants were randomly assigned to experimental conditions and were supervised as they completed the materials individually during class time.

2.1.2. Materials

Each participant read 1 of 4 different vignettes. The vignettes reflected a 2(Identification: low vs. high)×2(Constraint: medium vs. high) between-participants design. The initial section of each vignette was common to all four conditions. In this section two married couples, Susan and Bill and Elaine and Frank, are depicted on a Caribbean vacation, and subsequently on board an airliner returning home. It is revealed that Susan and Frank have been involved in a love affair and that Bill has discovered proof of the affair.

In the *High Identification* condition participants read:

> The humiliation and betrayal were almost more than he could bear. These were the two people he trusted most in the world. During the three days of the vacation that remained, he wrestled with the issue. He thought of many ways of retaliating. Finally, he decided that there was only one way he could deal with it. Bill decided that he would kill Frank.

In the *Low Identification* condition participants read:

> During the three days of the vacation that remained, he wrestled with the issue. Finally, he decided that if Susan and Frank wanted to be together, he would not stand in their way. He would confront them with the evidence and assume that whatever happened would be for the best. He really cared for both of them and wanted to be a forgiving person. He felt somewhat at peace with himself.

In the *High Constraint* condition, the next section of the narrative was the following:

> On the return trip home their plane was hijacked by a gang of 8 kidnappers. The pilot was forced to land in Bermuda, where the hijackers demanded a ransom of 5 billion dollars. To show the government their ruthlessness, the hijackers executed an elderly

UNCORRECTED PROOF

male passenger. They then seized two of the male passengers, Bill and Frank. The leader of the hijackers handed Bill a pistol with one bullet in it. With four machine guns pointed at him, Bill was ordered to shoot Frank in the head. He was told that if he did not obey, Frank, himself, and 10 other passengers would be killed. Bill realized that there was no way to resist or overpower the hijackers, because he and the other passengers were no match for 8 heavily armed men; any attempted heroics on his part would result in more loss of life than obeying the hijackers' orders.

In the *Moderate Constraint* condition, the same basic situation was depicted, but modified to lessen the degree to which Bill's shooting of Frank was compelled by the circumstances. The vignette in this condition altered the description to make resistance to and refusal of the hijackers' order appear a more viable behavioral option. It contained the following wording:

On the return trip home their plane was hijacked by a gang of 3 kidnappers. The pilot was forced to land in Bermuda, where the hijackers demanded a ransom of 5 billion dollars. To show the government their ruthlessness, the hijackers executed an elderly male passenger. They then seized two of the male passengers, Bill and Frank. The leader of the hijackers handed Bill a pistol with one bullet in it while another hijacker pointed a pistol at Bill. The third hijacker was in the pilot's cabin shouting angrily. At that moment, they were interrupted by an amplified voice ordering the hijackers to surrender immediately. Looking out the window, Bill saw that the plane was surrounded by heavily armed anti-terrorist forces. Bill quickly reviewed his options. He could try to persuade the hijackers that their situation was hopeless. He could stall until the anti-terrorist forces stormed the plane. The hijackers had been distracted by the arrival of the armed troops. Both the leader and the man holding a gun on Bill were nervous, frequently glancing out the windows of the plane. Perhaps, Bill thought, he could shoot the hijacker with the gun and the rest of the passengers could subdue the other two kidnappers. It was a risky move, but it could work. Bill thought he just might be able to pull it off, but the hijackers were angrily ordering him to "get on with it."

Next, in the *High Identification* condition, participants read:

Despite the desperate circumstances, Bill understood the situation. He had been presented with the opportunity to kill his wife's lover and get away with it. And at that moment Bill was certain about his feelings. He wanted to kill Frank. Feeling no reluctance, he placed the pistol at Frank's temple and proceeded to blow his friend's brains out.

Alternatively, in the *Low Identification* condition, participants read:

Bill was horrified. At that moment Bill was certain about his feelings. He did *not* want to kill Frank, even though Frank was his wife's lover. But although he was appalled by the situation and beside himself with distress, he reluctantly placed the pistol at Frank's temple and proceeded to blow his friend's brains out.

UNCORRECTED PROOF

After reading the vignettes, participants completed an 8-item questionnaire (Likert-type, 7-point scale).

2.1.3. Dependent variables

The study's principal dependent variable, the actor's degree of responsibility for his friend's death, was assessed by the first questionnaire item, "Bill is responsible for Frank's death." Seven additional, supplementary items assessed the participants' attitudes regarding the propriety of the actor and the action. These items were included as vehicles for exploratory analyses intended to shed light on the social cognition that is correlated with the attribution of moral responsibility. We had asked various philosophers and psychologists to suggest "What concepts are closely related to responsibility for an anti-social act?" These suggestions and our own intuitions were incorporated in the seven items, shown in Table 1. Because these items were to be examined with multivariate statistics, we limited their number.

2.2. Results

To test the principal hypothesis of the study, an initial two-way (Identification × Constraint) univariate analyses of variance (ANOVA) was performed upon the item worded, "Bill is responsible for Frank's death." This variable will be referred to hereafter as Bill's Responsibility. As hypothesized, significant main effects were found for Identification, $F(1,68)=6.83$, $P<0.02$, and for Constraint, $F(1,68)=5.02$, $P<0.03$. As depicted in Fig. 1, when Bill was identified with the act, he was judged more responsible than when not identified. Also Bill was judged more responsible when he operated under less constraint.

An exploratory factor analysis was performed on the seven supplementary questionnaire items. A principal-components analysis with varimax rotation was employed. Two factors emerged with eigen values greater than 1.0, accounting for 63.6% of the variance. Factor loadings of the items are shown in Table 1. An examination of the factor loadings for Factor 1 shows it to be constituted largely by items related to Bill's culpability and propriety. Factor 2 seems to be tapping the attribution of responsibility to sources other than Bill, namely the hijackers.

A factor scale corresponding to each factor was constructed by an unweighted summation the items that loaded higher than 0.50 on that factor. Four items were summed

Table 1
Factor loadings of dependent variables (Experiment 1)

Items	Factor 1	Factor 2
The hijackers are responsible for Frank's death.	−0.11	0.86
A person who does what Bill did should feel guilty.	0.87	0.05
Bill is a person of good character.	−0.39	0.26
The act Bill committed was wrong.	0.84	−0.08
Bill acted properly.	−0.64	0.13
Bill is to blame for Frank's death.	0.72	−0.17
The hijackers are to blame for Frank's death.	−0.05	0.89

UNCORRECTED PROOF

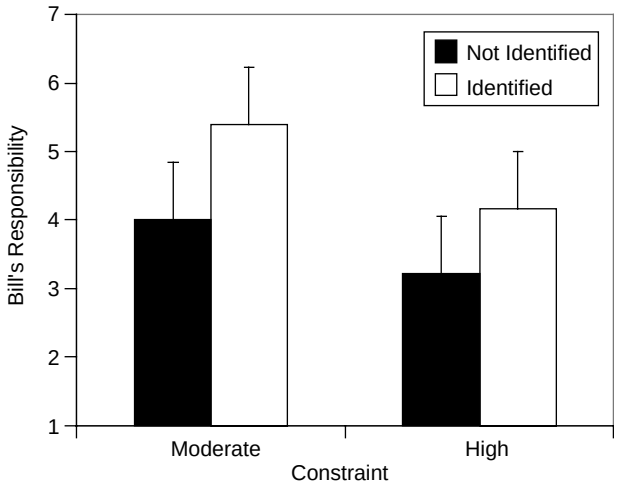

Fig. 1. Mean levels of responsibility attributed to actor different levels of identification constraint.

to produced the scale for Factor 1 and two items made up the scale for Factor 2 (values of items with negative loadings were reversed). The Factor 1 scale correlated highly with the study's primary dependent variable, Bill's Responsibility, $r(70)=0.78$, $P<0.001$. This scale was labeled Bill's Blameworthiness. The Factor 2 scale was labeled Hijacker Responsibility and was weakly correlated, negatively, with Bill's Responsibility, $r(70)=-0.11$, n.s. The scales for Factor 1 and Factor 2 were negatively correlated, $r(70)=-.24$, $P<0.05$.

Each of the two scales corresponding to Bill's Blameworthiness and Hijacker Responsibility was subjected to the same analysis employed on Bill's Responsibility, a two-way univariate ANOVA (Identification×Constraint). Given that conducting these additional exploratory ANOVAs raises the probability of making a Type I error within the experiment, we made a bonferroni-like adjustment by setting the statistical significance levels for the exploratory analyses at 0.0166. A significant main effect for Identification, $F(1,68)=11.74$, $P<0.001$, emerged on Bill's Blameworthiness. Higher Identification was associated with the higher scores on Bill's Blameworthiness, indicating the assignment of greater culpability and impropriety. Higher scores on Bill's Blameworthiness also were assigned when Bill operated under less Constraint, $F(1,68)=10.90$, $P<0.002$. Means associated with these effects are depicted in Fig. 2.

Scores on Hijacker Responsibility were significantly higher when Identification was lower, $F(1,68)=7.37$, $P<0.009$, reflecting that more responsibility was attributed to the hijackers when Bill was not identified with the shooting. A significant Identification× Constraint interaction emerged, $F(1,68)=13.78$, $P<0.0004$, showing that responsibility assigned to the hijackers was diminished only when Bill was both identified and operated under less constraint. The main effect for Constraint on Factor 2 failed to achieve statistical significance ($P>0.06$).

UNCORRECTED PROOF

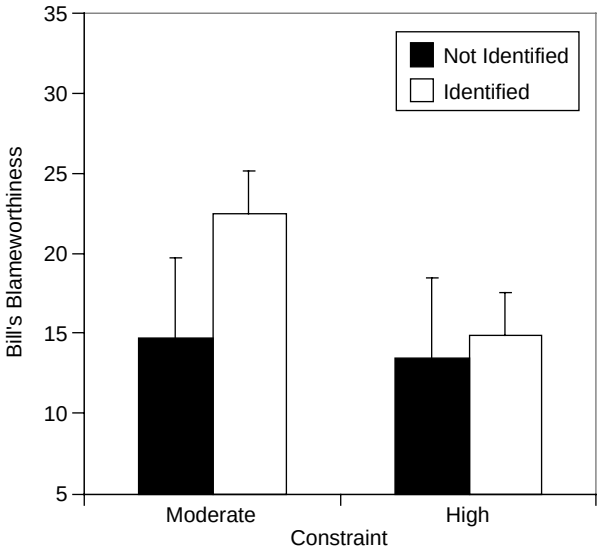

Fig. 2. Mean Factor 1 scores (Bill's Blameworthiness) at different levels of identification and constraint.

2.3. Discussion

Despite the presence of coercive circumstances that involved strong constraint upon behavior, the level of the actor's identification appeared to influence variables measuring the assignment of responsibility for an action. These findings, coupled with the fact that varying the level of constraint also affected responsibility attribution, led us to conduct a second experiment to replicate the effects observed and to examine the influence of identification upon responsibility attribution when the degree of constraint was increased to levels in excess of those employed in Experiment 1.

3. Experiment 2

3.1. Method

In Experiment 2 we sought extend the findings of Experiment 1 by further increasing the level of constraint under which the action was performed. The reader might be startled by the thought that multiple automatic weapons pointed at the actor would not constitute a maximally coercive situation. But some philosophers with whom we discussed our findings argued that, even in such circumstances, participants may have believed that it was possible for Bill to refrain from shooting Frank. These philosophers asserted that we had not produced a scenario in which the actor was *unable* to do otherwise, that we had not created a situation in which there was no alternate possibility.

To respond to this critique, we constructed a condition in which the actor operates under what we now term "absolute constraint," as contrasted with what we would now call

UNCORRECTED PROOF

the "practical constraint" operationalized in Experiment 1. In this Absolute Constraint condition Bill is administered a "compliance" drug that makes him unable to resist the demands of the hijackers. To insure the validity of this condition we instructed participants to "suspend disbelief" as to the facts specified, especially in regard to the power of a drug to effect total compliance. This is a familiar technique in philosophical "thought experiments" (see Doris & Stich, 2004) designed to test intuitions that cannot be readily evaluated using responses to plausible naturalistic scenarios. It is difficult to create a scenario in which an intentional action occurs, but is such that the actor indisputably could not have done otherwise, as indeed our philosophical interlocuters' comments on the previous experiment suggested. To this end, we asked participants to evaluate the behavior depicted in the vignette while assuming the absolute efficacy of the "compliance drug."

The design of the experiment was a 2×3 factorial. We examined the two levels of Identification: High and Low. The two Constraint conditions from Study 1 were replicated and a third condition, *Absolute Constraint*, was added.

3.1.1. Participants and procedure

Forty-eight (27 female and 21 male) University of California, Santa Cruz undergraduates enrolled in philosophy classes participated in the experiment as volunteers. Participants were randomly assigned to experimental conditions and were supervised as they completed the materials individually during class time.

3.1.2. Materials

Each participant read 1 of 6 different vignettes. The vignettes reflected a 2(Identification: low vs. high)×3(Constraint: moderate vs. high vs. absolute) between-participants design.

The initial section of each vignette was common to all six conditions and was identical to that employed in Experiment 1, describing the relationships among the principals. High and Low Identification were depicted using the exact wording of Experiment 1. In the Moderate and High Constraint conditions the identical wording from Experiment 1 was employed. The *Absolute Constraint* condition was created by the following wording:

> On the return trip home their plane was hijacked by a gang of 8 kidnappers…. They then seized two of the male passengers, Bill and Frank. The leader of the kidnappers injected Bill's arm with a "compliance drug"—a designer drug similar to sodium pentathol, "truth serum." This drug makes individuals unable to resist the demands of powerful authorities. Its effects are similar to the impact of expertly administered hypnosis; it results in total compliance. To test the effects of the drug, the leader of the kidnappers shouted at Bill to slap himself. To his amazement, Bill observed his own right hand administering an open-handed blow to his own left cheek, although he had no sense of having willed his hand to move. The leader then handed Bill a pistol with one bullet in it. Bill was ordered to shoot Frank in the head… But when Bill's hand and arm moved again, placing the pistol at his friend's temple, Bill had no feeling that he had moved his arm to point the gun; it felt as though the gun had moved itself into position. Bill thought he noticed his finger moving on the trigger, but could not feel any sensations of movement. While he was observing these events, feeling like a puppet,

UNCORRECTED PROOF

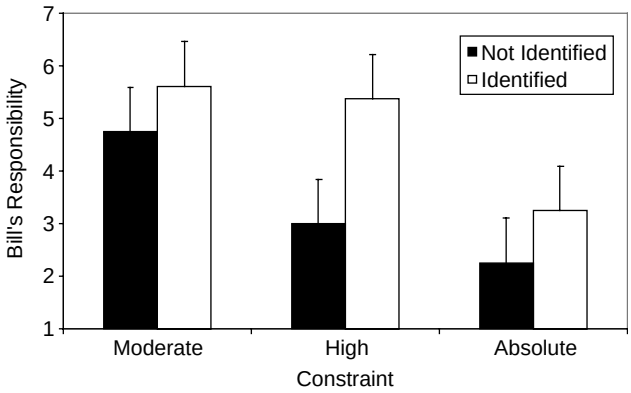

Fig. 3. Mean levels of responsibility attributed to actor at different levels of identification and constraint.

passively observing his body moving in space, his hand closed on the pistol, discharging it and blowing Frank's brains out.

3.2. Results

The same dependent variables used in Experiment 1 were employed in this experiment. Because our primary experimental hypothesis involved the assignment of responsibility, as in Experiment 1, an intitial two-way (Identification × Constraint) univariate analysis of variance was performed upon the item assessing Bill's Responsibility. Fig. 3 depicts the all the cell means of this analysis.[1] A significant main effect was found for Identification $F(1,42)=9.89$, $P<0.003$. When Bill was identified with the act, he was held more responsible than when not identified. A significant main effect for Constraint, $F(2,42)=9.87$, $P<0.0003$, indicated that when Constraint was higher, Bill was regarded as less responsible. A posteriori comparisons of the Constraint cell means, employing Tukey' s HSD, determined all pairwise differences between cell means to be statistically significant, except that between Moderate Constraint and High Constraint Ps <0.016.

As in Experiment 1, a principal-components analysis with varimax rotation was performed on the questionnaire item scores. Two factors emerged with eigen values greater than 1.0, accounting for 67.1% of the variance. The factor structure derived in Experiment 1 was cross-validated, in that a very similar pattern of item loadings was observed (see Table 2).

[1] The reader will note that even in the Absolute Constraint condition, when the actor also does not wish to commit the act, some responsibility is still attributed, suggesting something of a "floor effect" on the scale, i.e. that not all participants are willing to assign values at the extreme low end of the scale (see Table 4). Although a similar, but smaller, "ceiling effect" will emerge in Experiment 3, at the high end of the responsibility scale, we do not have data that tells us how much responsibility participants would assign to an actor who is both unconstrained in any apparent way and who also desires to commit the act.

UNCORRECTED PROOF

Table 2
Factor loadings of dependent variables (Experiment 2)

Items	Factor 1	Factor 2
The hijackers are responsible for Frank's death.	−0.22	0.90
A person who does what Bill did should feel guilty.	0.71	−0.17
Bill is a person of good character.	−0.36	0.22
The act Bill committed was wrong.	0.74	−0.01
Bill acted properly.	−0.73	0.12
Bill is to blame for Frank's death.	0.84	−0.14
The hijackers are to blame for Frank's death.	−0.12	0.94

The factors again were named Bill's Blameworthiness and Hijacker Responsibility. Creating a factor scale for each factor by an unweighted summation of those items that loaded at 0.50 or higher, produced two scales with identical item compositions to those of Experiment 1.

Each scale score for Bill's Blameworthiness and Hijacker. Responsibility was subjected to the same univariate ANOVA described above. A main effect for Identification on Bill's Blameworthiness emerged, $F(1,42)=7.95$, $P<0.008$. Bill's Blameworthiness was higher when the actor was identified than when not identified. A main effect also was observed for Constraint on Bill's Blameworthiness, $F(2,42)=4.79$, $P<0.014$. As depicted in Fig. 4, means on this variable were inversely proportional to the level of Constraint. A posteriori pairwise comparisons (Tukey's HSD) among Constraint cell means determined the only significant difference ($P<0.016$) to be that between the means of the Moderate Constraint and Absolute Constraint conditions.

Significant main effects on Hijacker Responsibility were observed for both Identification, $F(1,42)=14.76$, $P<0.0005$, and Constraint, $F(2,42)=14.41$, $P<0.0001$. Hijacker Responsibility was rated lower when the Bill was identified with the act. Hijacker Responsibility scores were directly proportional to the level of Constraint. When Bill operated under Absolute Constraint, Hijacker Responsibility scores were highest ($M=12.69$, $SD=1.92$), under High Constraint ($M=11.31$, $SD=2.09$) scores were

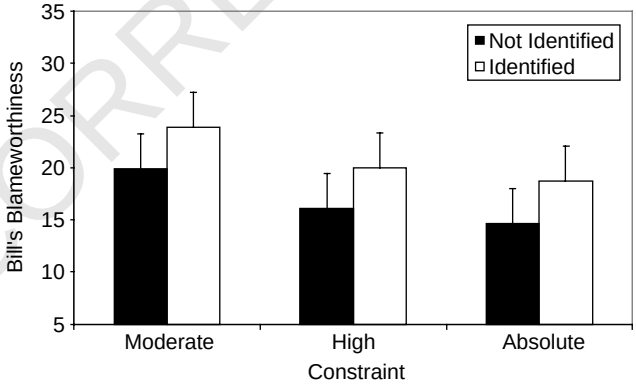

Fig. 4. Mean factor 1 scores (Bill's Blameworthiness) at different levels of identification and constraint.

UNCORRECTED PROOF

intermediate, and Moderate Constraint ($M = 8.69$, SD = 3.48) yielded the lowest scores on Hijacker Responsibility. A posteriori comparisons of every Constraint cell mean employing Tukey' s HSD determined all pairwise differences between cell means to be statistically significant except that between Moderate Constraint and High Constraint, Ps < 0.016. As in Experiment 1, a significant Identification × Constraint interaction occurred, $F(2,42) = 4.58$, $P < 0.016$, reflecting the disproportionately low assignment of responsibility to the hijackers when Bill was both identified and operating under the least constraint.

3.3. Discussion

The findings of Experiment 2 replicated and extended the principal findings of Experiment 1. Our data suggested that observers are inclined, under some circumstances, to hold actors who identify with an act more responsible than unidentified actors, even when it is highly plausible to suppose that the actor "could not have done otherwise." These findings are discussed at length in the General Discussion.

4. Experiment 3

In the previous two experiments, we created stories designed to cause experimental participants to perceive that the act of killing another person occurred under considerable constraint. We wrote different versions of the experimental scenarios designed to differ in the degree of constraint they depicted. Although many of our actor responsibility measures are indirect indications of the degree of constraint that the participants perceived, we did not directly ask participants in Experiments 1 and 2 to rate the degree to which the actor had been constrained, because we did not want to alert them to our interest in theoretical issues having to do with levels of constraint. So in the present study, we present data on the degree of constraint or coercion to which the actor was perceived to be subjected in the experimental vignettes.

We presented participants with three vignettes (with the identification manipulation removed), each containing one of the three levels of constraint used in the previous experiments. We also included a condition in which the actor shoots the victim under *no* ostensible constraint or coercion.

4.1. Method

4.1.1. Participants and procedure

Forty-eight (28 female and 20 male) Princeton University students participated in the experiment in order to fulfill requirements for participation in Psychology Department research. Participants filled out various short "paper and pencil" research projects from various investigators. The projects were chosen for their theoretical and mundane dissimilarity, and the order of the questionnaires randomized to minimize order effects. Participants reported to a specified room at set times, and were given packets containing the various questionnaires by an experimenter.

UNCORRECTED PROOF

4.1.2. Materials

Each participant read all four vignettes. The No Constraint vignette contained the material common to all vignettes and included the material below:

> Bill was shocked and upset. After wrestling with the issue, he purchased a black market handgun. On the last day of their vacation, Bill followed Frank as he walked through the back streets of the city they were visiting. When the opportunity presented itself, he placed the pistol at Frank's temple and proceeded to blow his friend's brains out.

The No Constraint vignette was presented first to all participants. The order of the remaining three vignettes was counterbalanced to yield six different orders of presentation.

Participants rated the conduct depicted in each vignette on a 6-item, manipulation-check questionnaire that assessed the degree to which Bill "was constrained," "was forced," "was free to do other than he did," "had a choice," "could have behaved differently," and whether it was "reasonable" to expect him to have behaved differently. Items were Likert-type, with 7-point scales.

4.2. Results

Because the six questionnaire items were highly correlated ($rs > \pm 0.28$), instead of univariate tests a one-way repeated-measures multivariate analysis of variance was conducted. Constraint was the classification variable (4 levels, within participants) and all six dependent variables were included. This analysis was significant, $F(3, 45) = 153.77$, $P < 0.001$. A posteriori Tukey's HSD tests indicated that each level of Constraint was significantly different from every other level, $Ps < 0.005$. No Constraint was associated with the greatest perceived freedom, Moderate Constraint linked with less freedom, High Constraint with still less freedom, and Absolute Constraint associated with the lowest ratings of freedom. Table 3, which shows the means for Item 3 ("was free to do other than he did"), provides some sense of the relative degree of freedom that was attributed to Bill in the four conditions.

After reading the No Constraint vignette, participants also were asked to rate Bill's responsibility for Frank's death on the same item that was the principal dependent variable in Experiments 1 and 2. The mean rating of Bill's Responsibility ($M = 6.44$, SD $= 1.38$) in the No Constraint condition was higher than any cell mean rating on that item for any condition in Experiments 1 or 2. A somewhat more concrete understanding of the effect of the constraint manipulation in Experiments 1 and 2 upon responsibility ascriptions to the actor can be achieved by comparing the mean for Bill's Responsibility under No Constraint in Experiment 3 with the means associated with the different levels of Constraint in Experiments 1 and 2. This is done in Table 4.

Table 3
Mean assignment of freedom to the actor (Bill) by level of situational constraint (7-point scale)

Item	No constraint	Moderate constraint	High constraint	Drug-induced, "absolute" constraint
Free "to do other than he did"	6.63	5.21	3.08	1.98

UNCORRECTED PROOF

Table 4
Mean assignment of responsibility to the actor (Bill) by level of situational constraint and level of identification (7-point scale)

Actor's level of identification	Moderate constraint	High constraint	Drug-induced, "absolute" constraint	No constraint (identification not manipulated)
Identified	5.49	4.81	3.25	6.44
Not identified	4.38	3.11	2.25	

Means for moderate constraint and high constraint conditions are averaged across Experiments 1 and 2.

4.3. Discussion

The results of Experiment 3 provided support for the internal validities of Experiments 1 and 2. It would appear that our manipulation was successful, in that the degree of coercion in the circumstances we portrayed caused participants to adjust their ratings of the actor's freedom and responsibility accordingly.

5. General discussion

In the preceding experiments we attempted to elucidate some aspects of lay moral cognition. We found participants' attributions of responsibility for an action to be influenced by the actor's attitude toward that action, even when the action was causally constrained to such a degree that there were no other behavioral options. The degree to which actors "identified" with an action was strongly associated with responsibility for the action being assigned to them.

Our results have relevance to various discussions of responsibility attribution in the psychology literature. Attributional analyses employing a "discounting principle" would require assignments of moral responsibility to vary largely with the extent to which the actor exerts causal control on the outcome. But our data suggest that the attribution of moral responsibility takes into account non-causal elements, such as identification, in addition to causal factors. These data thus augment recent empirical work that emphasizes observers' perceptions of the desires and intended outcomes of the actor (Alicke, 2000; Schlenker et al. 1994; Weiner, 1995). Our data support the view that information about outcomes that an actor desires can moderate or override the attributional effects of the actor's perceived control over events.

Schlenker et al. (1994) describe a component of responsibility attribution they call the "identity-event linkage," which is the degree to which an actor is perceived as linked with an action, a link that has to do with intending to bring the action about or acting to bring it about, and doing so in the absence of excusing circumstances. Our findings are consistent with much of this formulation, but we would suggest a refinement of the authors' view that "personal control over the event" is required. In the present studies an actor who desired an outcome was judged to be to some extent responsible, even though there were extenuating circumstances of the most extreme sort. Here participants appear either to have perceived a link between the actor and the public event, based largely on the actor's attitudes, or to

UNCORRECTED PROOF

have assigned responsibility for the action based upon what participants construed to be an internal and private event over which the actor did have personal control: his internally desiring the outcome of an act that he was coerced to produce.

Alicke's (2000) theory of culpable control has deep roots in philosophy and therefore addresses issues relevant to our findings. Alicke treats intention as separate from causation in his analysis of his core construct of "personal control." His work suggests that the perceived linkages among a person's intentions, behavior, and the ultimate outcomes of that behavior are the key factors in judgments of culpability. Our work harmonizes with his approach, in that we can be viewed as exploring situations in which there is either consistency or inconsistency between an actor's desires and the actor's behavior, as well as circumstances in which there is either congruence or incongruence between desire and outcome. Alicke also stipulates that the same behavior can be judged as more or less culpable simply on the basis of the actor's relevant attitudes. Our case to be accounted for is the one in which some responsibility is attributed to an individual even when he was coerced to take an action, because the outcome of that action was one he desired. In Alicke's theory the machinery to account for this case is available in his constructs of "volitional behavior control" and "volitional outcome control." As do most models of responsibility attribution, his theory assigns an important role to the mitigating influence of external constraints in reducing the assignment of blame. This formulation is consistent with our general finding of reduced responsibility assigned for antisocial behavior as external coercion to commit that behavior increased. What is not clear is whether our finding of effects for identification in the presence of overwhelming constraint would be predicted by his theory. Some indication that the theory might allow for such phenomena is Alicke's assertion that observers tend to assume agency and assign blame as a default, subsequently adjusting attributions for mitigating factors. This then would be the classic anchor and adjustment process, in which it is known that the adjustment is generally insufficient to move the attributions appropriately far away from the initial anchor.

With respect to the philosophical literature, it appears that in at least some contexts the tacit theory of responsibility employed by social perceivers is not straightforwardly incompatibilist; in particular, actors may be held responsible even in circumstances where it is apparent they were coercively constrained to act as they did. For philosophers who believe theories of responsibility should be strongly informed by everyday social practice, as P.F. Strawson (1982) advocated, the data suggest that compatibilist theories may not be contrary to lay practices. Our data suggest that the assumption, made by many philosophers, that lay perceivers are "natural incompatiblists" (Kane, 2003, p. 300) is open to question.

This is not to argue that folk theories of responsibility are uniformly compatiblist. In fact, we would suggest that folk theories of responsibility are most likely *contextualist*, meaning that differing considerations are salient to moral responsibility attribution in different contexts, and that patterns of responsibility attribution may also vary culturally and developmentally. For example, recent empirical work by Nichols (in press) suggests that, at least in some situations, children (ages 3–6) treat the "could have done otherwise" condition of the principle of alternate possibilities as a precondition of human agency. The discourse of responsibility takes place in widely varying contexts and operates in connection with diverse human interests. We might suspect that the along broad spectrum

UNCORRECTED PROOF

of moral cognition related to responsibility ascription, which ranges from the determination of criminal liability, to the assignment of credit for scientific discovery, to deciding which sibling should have to clean up the spilled milk, complex considerations often come into play. How "responsibility" is assessed in varying contexts is a wonderful question for future research.

The present research supports some preliminary conclusions. We have provided evidence that lay attributions of responsibility may, in some circumstances, accord with philosophical views in which freedom and determinism are regarded to be compatible. We have shown also that the ascription of responsibility is an even more complex process than many theorists previously have contended. Adding the concept of identification to the other factors known to influence responsibility attribution may assist in understanding these complexities. Our findings indicate that further explorations of the construct of identification are warranted.

One question worthy of further exploration concerns what might be called the duration and depth of identification the actor has with the eventual outcome. All of us can have fleeting desires to produce some morally negative outcome, such as seriously harming an individual who criticizes a manuscript of ours. If at that moment some external constraints cause us to injure the critic, do observers attribute the level of responsibility we have observed in the present data (where the intention to harm is more enduring) for this transitory intention to harm the other? Or do the observers treat the fleeting, emotion-produced intentions of the moment as not counting as "identifying" with coerced harms of the other? Questions such as these invite research attention.

The concept of identification, as we employ it, is somewhat broader than but seems to include elements of "intention" or "intentional behavior," as developed in the work of Malle, Knobe, & colleagues (e.g., Malle, 1999; Malle & Knobe, 1997a,b). The concept of identification also is related to that of "metadesire" or "second-order" desire, a desire to have certain desires, and to have those desires influence one's behavior. Recent research by Pizzaro, Uhlmann, & Salovey (2003) found that manipulating perceptions of actors' metadesires affected assignments of praise and blame for morally relevant conduct, a result consistent with our data. Future research should be directed toward exploring the connections between identification and such related concepts.

In summary, even in the case when an act was committed under conditions of absolute and overwhelming constraint, responsibility attributions were powerfully affected in the predicted direction by the identification manipulation. This finding is in violation of a venerable psychological principle of social cognition, the "discounting principle," and an equally venerable philosophical principle, the "principle of alternative possibilities." In our estimation the best explanation of this phenomenon is that the cognition involving moral attribution is strongly influenced by extra-causal factors, i.e. factors other that those that are likely to figure in the most careful and thoughtful *causal* explanation of the behavior in question. One such factor is the evaluative attitudes of the actor; what it is that the actor wants to come about or wants not to come about. As we remarked above, moral cognition involves evaluative as well as causal dimensions; it is an activity broader in scope than the activity ascribed to "the intuitive psychologist" familiar in the social cognition literature. What causes people to attribute responsibility, to praise or blame, is to

UNCORRECTED PROOF

some extent what is believed to be in the "heart" of the actor and this is so even for actions committed under overwhelmingly coercive or constraining circumstances.

6. Uncited references

Ekstrom (2000); Heider (1958); McKenna (2001).

Acknowledgements

The authors wish to express their gratitude to Lesley Allen, Joel Cooper, and Michael Gara for various forms of assistance in this research and to thank the editor and three anonymous reviewers for valuable comments on the manuscript.

References

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, *126*, 556–574.

Bratman, M. E. (1996). Identification, decision, and treating as a reason. *Philosophical Topics*, *24*, 1–18.

Darley, J. M., & Cooper, J. (Eds.). (1998). *Attribution and social interaction: The legacy of Edward E. Jones*. Washington, DC: American Psychological Association.

Darley, J. M., & Shultz, T. R. (1990). Moral rules: Their content and acquisition. *Annual Review of Psychology*, *41*, 525–556.

Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. Cambridge: Cambridge University Press.

Doris, J. M., & Stich, S. P. (2004). As a matter of fact: Empirical perspectives on ethics. In F. Jackson, & M. Smith (Eds.), *The Oxford handbook of contemporary analytic philosophy*. Oxford: Oxford University Press in press.

Ekstrom, L. W. (2000). *Free Will: A philosophical study*. Boulder: Westview Press.

Fein, S., Hilton, J. L., & Miller, D. T. (1990). Suspicion of ulterior motivation and the correspondence bias. *Journal of Personality and Social Psychology*, *58*, 753–764.

Frankfurt, H. G. (1988). *The importance of what we care about: Philosophical essays*. Cambridge: Cambridge University Press.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Hilton, J. L., Fein, S., & Miller, D. T. (1993). Suspicion and dispositional inference. *Personality and Social Psychology Bulletin*, *19*, 501–512.

Jones, E. E. (1990). *Interpersonal perception*. New York: W.H. Freeman and Company.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz, *Advances in experimental social psychology* (Vol. 2) (pp. 219–266). New York: Academic Press, 219–266.

Kane, R. (2002). Introduction: The contours of contemporary free will debates. In R. Kane (Ed.), *The Oxford handbook of free will* (pp. 3–41). New York: Oxford University Press, 3–41.

Kane, R. (2003). Responsibility, luck, and chance: Reflections on free will and determinism. In G. Watson (Ed.), *Free Will*2nd ed. (pp. 299–321). New York: Oxford University Press, 299–321.

Kelley, H. H. (1972). Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1–26). Morristown, NJ: General Learning Press, 1–26.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review*, *3*, 23–48.

UNCORRECTED PROOF

Malle, B. F., & Knobe, J. (1997a). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*, 101–121.

Malle, B. F., & Knobe, J. (1997b). Which behaviors do people explain? A basic actor–observer asymmetry. *Journal of Personality and Social Psychology*, *72*, 288–304.

McClure, J. (1998). Discounting causes of behavior: Are two reasons better than one? *Journal of Personality and Social Psychology*, *74*, 7–20.

McKenna, M. (2001). Source incompatibilism, ultimacy, and the transfer of non-responsibility. *American Philosophical Quarterly*, *38*, 37–51.

Nagel, T. (1986). *The view from nowhere*. New York: Oxford University Press.

Nichols, S. (in press). The folk psychology of free will: Fits and starts. *Mind & Language*.

Pizzaro, D., Uhlmann, E., & Salovey, P. (2003). Asymmetry in judgments of moral praise and blame: The role of metadesires. *Psychological Science*, *14*, 267–272.

Quattrone, G. A. (1982). Behavioral consequences of attributional bias. *Social Cognition*, *1*, 358–378.

Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology*, *83*, 789–803.

Robinson, D. R. (1996). *Wild beasts and idle humours: The insanity defense from antiquity to the present*. Cambridge: Harvard University Press.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz, *Advances in experimental social psychology* (Vol. 10) (pp. 174–221). New York: Academic Press, 174–221.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation: Perspectives of social psychology*. Philadelphia, PA: Temple University Press.

Schlenker, B. R., Britt, T. W., Pennington, J., Murphy, R., & Doherty, K. (1994). The triangle model of responsibility. *Psychological Review*, *101*, 632–652.

Shaver, K. G. (1985). *The attribution of blame: causality, responsibility, and blameworthiness*. New York: Springer.

Shultz, T. R., & Schleifer, M. (1983). Towards a refinement of attribution concepts. In J. Jaspars, F. Fincham, & M. Hewstone (Eds.), *Attribution theory and research: Conceptual, developmental, and social dimensions* (pp. 37–62). London: Academic Press, 37–62.

Smith, J. W. (1961). Impossibility and morals. *Mind*, *70*, 362–375.

Strawson, P. F. (1982). Freedom and resentment. In G. Watson (Ed.), *Free will* (pp. 59–80). New York: Oxford University Press, 59–80.

Velleman, J. D. (1992). What happens when someone acts? *Mind*, *101*, 461–481.

Wallace, R. J. (1994). *Responsibility and the moral sentiments*. Cambridge, MA: Harvard University Press.

Watson, G. (1996). Two faces of responsibility. *Philosophical Topics*, *24*, 227–248.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. New York: Guilford Press.

UNCORRECTED PROOF