

Situationism and Agency

ALFRED R. MELE

Department of Philosophy, Florida State University

JOSHUA SHEPHERD

Department of Philosophy, Florida State University

ABSTRACT

Research in psychology indicates that situations powerfully impact human behavior. Often, it seems, features of situations drive our behavior even when we remain unaware of these features or their influence. One response to this research is pessimism about human agency: human agents have little conscious control over their own behavior, and little insight into why they do what they do. In this paper we review classic and more recent studies indicating “the power of the situation,” and argue for a more optimistic response. In our view, though psychological research indicates situational influence, it also indicates that knowledge about the impact of situations on behavior can boost agents’ power to counteract harmful situational effects.



Any system of practical ethics makes some presuppositions about human agents. It may be assumed, for example, that many people are capable of making informed, conscious decisions about what to do in a wide range of situations and are capable of executing many such decisions. If people lack these capacities, the many practical injunctions that flow from ethical discourse come to seem misguided.

The assumption that people are capable of making informed, conscious decisions has been called into question by work in the human sciences. To take just one example, Benjamin Libet has argued, on the basis of his well-known neuroscientific findings, that, with one kind of exception, we never consciously decide what to do

(1985, 1999, 2004). The exception is for cases in which we become conscious of an urge or intention to do something. In some such cases, Libet claims, we are capable of consciously vetoing the urge or intention (1985, 1999, 2004).

One of us has argued that the neuroscientific case Libet presents for his thesis is far from persuasive (Mele 2009). Here we take up a related challenge to the common assumption mentioned in our opening paragraph and some related assumptions. The primary challenge we explore comes from the situationist (or “situationalist”) literature in psychology. This literature has received significant attention from philosophers in connection with moral virtue and moral character. Some see it as posing a serious challenge to the claim that there is such a thing as moral character (Doris 2002, Harman 1999), and others disagree (Kamtekar 2004, Sabini and Silver 2005, Sreenivasan 2002). We take a different approach here. Section One provides some background on situationism and briefly describes three well-known experiments on it. In section two, we contrast two perspectives on this literature, one pessimistic and another optimistic, and we side with the optimists. In our view, knowledge about the influences of situations on behavior can boost agents’ power to counteract harmful situational effects. In section three, we review recent work indicating that social cues also strongly influence behavior in a way that is mediated by unconscious processes. Again, we offer an optimistic interpretation of the import of this work for our capacity to control our behavior, and we provide some empirical support for our interpretation. In section four, we discuss some additional evidence that supports our optimism about agential control. The upshot is that although unconscious processing certainly influences behavior, there is reason to think that education about the influence of pertinent stimuli can play an ameliorating role in cases of undesirable influence. Section Five wraps things up.

SECTION ONE: SOME CLASSIC SITUATIONIST FINDINGS

Matthew Lieberman writes:

If a social psychologist was going to be marooned on a deserted island and could only take one principle of social psychology with him it would undoubtedly be “the power of the situation.” All of the most classic studies in the early days of social psychology demonstrated that situations can exert a powerful force over the actions of individuals. . . . If the power of the situation is the first principle of social psychology,

a second principle is that people are largely unaware of the influence of situations on behavior, whether it is their own or someone else's behavior. (2005, p. 746)

In the present section we describe three of these classic studies.

In a study by John Darley and Bibb Latané (1968), participants were led to believe that they would be talking about personal problems associated with being a college student. Each was in a room alone, thinking that he or she was talking to other participants over a microphone. Sometimes participants were led to believe that there was only one other participant (group A), sometimes that there were two others (group B), and sometimes that there were five others (group C). In fact, the voices the participants heard were recordings. Participants were told that while one person was talking, the microphone arrangement would not let anyone else talk. At some point, the participant would hear a person—the “victim”—say that he felt like he was about to have a seizure. The victim asks for help, rambles a bit, says he is afraid he might die, and so on. His voice is abruptly cut off after he talks for 125 seconds, just after he makes choking sounds. The percentage figures for participants who left the cubicle to help before the voice was cut off are as follows: group A 85%, group B 62%, group C 31%. Also, all the participants in group A eventually reported the emergency, whereas only 62% of the participants in group C did this.

Clearly, participants' beliefs about how other many other people could hear the voice—none, one, or four—had an effect on their behavior. Even so, there being one or four other people around to help the victim seems not to be a reason not to help.

Philip Zimbardo's Stanford prison experiment (Haney et al. 1973, Zimbardo n.d., Zimbardo et al. 1973) began with newspaper ads for male college students willing to take part in an experiment on prison life. The volunteers selected as prisoners were arrested at their residences, handcuffed, searched, and driven in a police car to a Palo Alto police station. From there, after being fingerprinted and placed in a detention cell, they were driven to a mock prison built in the basement of the Stanford psychology building. When they arrived, they were stripped and sprayed with deodorant. Then, after being given a prison uniform and photographed, they were locked in cells. There were three small cells—six by nine feet—for the ten prisoners and a very small solitary confinement cell. There were also rooms for volunteers selected as guards. Much of the activity was videoed by hidden cameras. Concealed microphones picked up conversations.

The plan was to run the experiment for two weeks. The prisoners were there twenty-four hours per day. The guards worked eight-hour shifts and then went home. Prisoners had three simple, bland meals a day and the same number of supervised toilet visits. They were also lined up three times each day to be counted and were always referred to by a number worn on their uniform—never by their name. They had two hours of free time each day to write letters or read—unless that privilege was taken away. And they had chores to do—cleaning toilets and the like. It is interesting that during their free time, 90% of what the prisoners talked about had to do with their prison life.

Zimbardo and coauthors report that “five prisoners had to be released because of extreme emotional depression, crying, rage and acute anxiety” (Haney et al. 1973, p. 81). Although the experiment was supposed to last two weeks, Zimbardo ended it after just six days. One prisoner had to be released after thirty-six hours owing to “extreme depression, disorganized thinking, uncontrollable crying and fits of rage” (Zimbardo et al. 1973). Another developed a psychosomatic rash.

Several of the guards became bullies, and those who did not participate in the bullying allowed it to continue. The harassment increased each day. Counting of prisoners, which originally took ten minutes, sometimes went on for hours. During these counts, prisoners were encouraged to belittle each other. Over time, the prisoners’ attitude toward one another reflected the guards’ attitude toward them. Insults and threats escalated, and so did commands to do pointless or demeaning tasks. Guards sometimes made prisoners clean toilets with their bare hands.

Pointless tasks included moving boxes back and forth from one closet to another and picking thorns out of blankets after guards had dragged the blankets through bushes. Sometimes prisoners would be made to do push ups while guards stepped on them. Guards would wake prisoners up in the middle of the night to count them. Sometimes they would deny them their scheduled leisure time just for the fun of it, or lock them in a solitary confinement cell for no good reason—a seven foot tall broom closet two feet wide and two feet deep. After the 10:00 p.m. lockup, prisoners often had to use buckets in their cells as toilets. On the second day of the experiment, prisoners staged a protest. The guards used a fire extinguisher to spray them, stripped them, and put the leaders in solitary confinement.

The guards created a privilege cell to sow dissension among the prisoners. The good prisoners would use the cell and get better treatment, including better food. After a while, to confuse the prisoners, the ones who seemed worse got the

privileges. Some of the guards became sadistic; and, of course, Zimbardo was as interested in the effects on the guards as the effects on the prisoners.

Bad effects of the situation showed up both in prisoners and in guards. The guards fell into three types. Some were tough but fair, some were good guys who did small favors for prisoners, and about a third were hostile and abusive. None of the testing the experimenters did in advance predicted which of the students would become power-loving guards. Some of the guards were disappointed that the experiment ended early; they enjoyed their power.

One of the prisoners felt sick and wanted to be released. He cried hysterically while talking with Zimbardo, in his role as prison superintendent, and a priest. After Zimbardo left the room to get the prisoner some food, the other prisoners began to chant that this one was a bad prisoner. When Zimbardo realized that the prisoner could hear this, he ran back into the room. He writes:

I suggested we leave, but he refused. Through his tears, he said he could not leave because the others had labeled him a bad prisoner. Even though he was feeling sick, he wanted to go back and prove he was not a bad prisoner. At that point I said, 'Listen, you are not #819. You are [his name], and my name is Dr. Zimbardo. I am a psychologist, not a prison superintendent, and this is not a real prison. This is just an experiment, and those are students, not prisoners, just like you. Let's go.' He stopped crying suddenly, looked up at me like a small child awakened from a nightmare, and replied, 'Okay, let's go'. (Zimbardo n.d.)

This episode makes especially salient how deeply participants were drawn into their roles.

We turn to Stanley Milgram's famous studies of obedience, beginning with the study first reported in Milgram 1963. Participants were forty men between the ages of twenty and fifty and from many different walks of life. The cover story was that the experimenter was conducting an experiment on how punishment is related to memory.

The participant meets the experimenter and a confederate of his. The participant is told that he and the other man will draw slips of paper from a hat to see which of them will be the "teacher" and which the "learner." In fact, the participant is always the teacher. He hears the cover story and sees where the learner will sit during the experiment—in a chair in which the learner will supposedly receive an electric

shock from the teacher each time he gives an incorrect answer. The teacher watches the learner being strapped into the electric chair and is told the straps will prevent him from moving too much when he is being shocked. The teacher then moves to another room where he can no longer see the learner. Milgram reports that, with a few exceptions, participants believed the setup was real.

Participants are shown an array of thirty levers, each associated with different degrees of shock. The lowest shock is for the first incorrect answer, the second lowest is for the second wrong answer, and so on. Sets of levers—mainly sets of four—are labeled. About half way through, the label is “intense shock,” followed by “extreme intensity shock,” “danger: severe shock,” and finally “XXX.”

The learner answers by pressing a button. At one point during the experiment—after he has received his twentieth shock—the learner pounds on the wall, and from then on he does not answer any more questions. The twentieth shock was delivered by the fourth lever in the “intense shock” level. The shock levels were also labeled with voltage numbers. This one was 300 volts. Before shocking the learner, the teacher had to report the voltage of the shock he was about to administer: 15 at the beginning all the way up to 450 at the end. At the beginning of the experiment, the scientist told the teacher that “Although the shocks can be extremely painful, they cause no permanent tissue damage.” When participants raised the issue of stopping the experiment, they were given stock replies ranging from “Please continue” to “You have no other choice, you must go on.” The scientist started with a simple request to continue and eventually moved up to the “no choice” response if the participant persisted in talking about stopping.

Twenty-six of the forty participants continued shocking all the way to the end. (Teachers were told that no answer counted as a wrong answer.) No participant stopped shocking before the twentieth shock. Five stopped right after that one. Four stopped after the next one: it was the first shock in the series labeled “extreme intensity shock” and the first shock in response to a non-answer. The other four dropped out a bit later.

Milgram reports that the participants displayed enormous tension, fits of nervous laughter, twitching, stuttering, sweating, and the like. And when they talked about stopping, a calm reply by the experimenter often worked: “The experiment requires that you continue,” “It is absolutely essential that you continue,” or the like. If a participant refused to continue after being told he had no choice, the experiment

was terminated and the participant was debriefed. This “no choice” response was the last in a series of four stock responses by the experimenter.

Milgram conducted many versions of this experiment. A brief description of three additional versions will prove useful. In Voice-Feedback (Milgram 1974, experiment 2), the teacher could now hear the learner speak. The learner grunts in response to the 75 volt shock and the slightly later ones. At 120 volts—labeled moderate—he shouts and says the shocks are becoming painful. He groans after the next shock, and refuses to continue in response to the one after that—the tenth shock. This goes on with increasing intensity for several more shocks. At 180 volts, the learner screams that he cannot stand the pain. By 270 volts he is screaming in agony. At 300 volts—the twentieth shock—he desperately shouts that he will not provide any more answers. And he repeats this after the next shock—after emitting a violent scream. After all subsequent shocks, he shrieks in agony. Twenty-five of the forty participants shocked all the way to the end.

In two other versions of the experiment, the teacher was brought much closer to the learner, but everything else was very similar—the groaning, screaming, and so on. There were forty participants in each. In one version (Proximity: Milgram 1974, experiment 3) the teacher was just a foot and a half from the learner and could see him clearly. In the other (Touch-Proximity: Milgram 1974, experiment 4), the learner could remove his hand from a shock plate in order to avoid being shocked, and the teacher would have to force the learner’s hand onto the plate in order to shock him. There were forty participants in each. In Proximity, sixteen participants continued to the end. In Proximity-Touch, twelve did.

SECTION TWO: PERSPECTIVES ON THE CLASSIC STUDIES

The findings we described (along with findings of many related studies) certainly are interesting. What should we make of them? According to a pessimistic view, they suggest that people have very little control over their behavior—that human behavior is largely driven by the situations in which people find themselves and the effects these situations have on automatic behavior-producing processes.

We are not so pessimistic. A few days after the tragic events of September 11, 2001, a friend said “That will never happen again.” He explained that, in his view, people would learn from what happened, and, henceforth, a plane full of passengers would not go down without a fight. They would resist, and they would overpower

their foes. That was an uplifting thought (a thought inspired in part by news reports that passengers and crew on United Airlines flight 93 had heard about the earlier crashes and attempted to regain control of the plane).

The role of “passenger” on a commercial flight is pretty well defined. Passengers are to sit down, fasten their seat belts, keep them fastened until they are told they are permitted to get up, refrain from being disruptive in any way, and, in general, obey the airline employees. For the most part, if there is a disturbance, passengers expect the flight crew to deal with it. The passengers’ situation involves ingredients of the three studies we described. The prisoners and guards occupy a role in Zimbardo’s studies; so do passengers. Obedience to pertinent authority figures is something typical passengers share with typical participants in Milgram’s studies. And when there is a disturbance on a plane, nonintervention by passengers is not surprising, especially given that such disturbances are matters the airline employees are expected to handle. In the bystander study that we described, participants had no reason to believe that an authority figure (the experimenter) was aware of the apparent emergency. So nonintervention by airplane passengers would seem to be more predictable, other things being equal.

Now, if behavior is driven by situations in such a way that new, consciously processed information is out of the behavior-producing loop, then our friend was way too optimistic. But we are inclined to agree with him. If we had had the horrible misfortune to be on one of the airliners that hit the World Trade Center years ago, we probably would have refrained from intervening and hoped that the airline employees would handle things. In light of what we learned, we predict that our reactions would be different now. Our expectation is that if a passenger or two attempted to intervene, others would join in.

This last remark is a window on our optimism. Behavioral education starts at an early age. Parents try to teach their toddlers to control potentially harmful impulses, and they enjoy a considerable measure of success. Parents also teach respect for parental authority; and they engage in moral education, which also involves instruction in self-control. Of course, parents can only teach what they are familiar with. And a lot more is known now about factors that influence human behavior than was known fifty years ago. Our view is that this knowledge should be put to good use, and not only in child rearing.

One often sees articles for public consumption claiming that neuroscientists have shown that free will is an illusion. One of us has made various attempts to

debunk such claims (see, e.g., Mele 2009 and Mele n.d.), but our point now is that lots of people find striking “news” about human behavior interesting. The classic studies that we described are not news now, of course; but they continue to be cited in new studies on situationism or automaticity. One way to spin news about these stories is pessimistic: for example, being in a group that witnesses an emergency has an enormous effect on your behavior, and there is nothing you can do about it. Another is not: and now that you know about the bystander effect, what will you do should you find yourself in a group that witnesses an emergency?

There are plenty of self-help books on self-control. People learn techniques for resisting or avoiding temptation with a view to making their lives go better. People who read such books know what they want to avoid—binge eating, gambling, binge drinking, or whatever—and they try to learn how to avoid it. When a cause of harmful behavior flies under everyone’s radar, not much can be done about it. But once a cause of harmful action or inaction is brought to light, prospects for amelioration may become brighter.

A public that is educated about the bystander effect is less likely to display it. The same is true of undue or excessive obedience to authority. In the latter sphere, matters are delicate. Obedience to authority is important for civil society. Because it is useful, it is instilled by parents, teachers, and so on; and it tends to become habitual in many people. But we also know the evils to which it can lead. Milgram’s work was motivated partly by a desire to understand how ordinary German citizens who became rank and file military personnel ended up committing atrocities. Obedience to authority is an important part of his answer. It would seem that the socialization of obedience to authority should include education about proper limits to obedience. Milgram writes: “In growing up, the normal individual has learned to check the expression of aggressive impulses. But the culture has failed, almost entirely, in inculcating internal controls on actions that have their origin in authority. For this reason, the latter constitutes a far greater danger to human survival” (1974, p. 147). Education can lessen this danger.

What about Zimbardo’s findings? They have obvious implications for the training of prison guards, and the implications clearly extend to people whose jobs give them considerable power over others—police, for example. But the import of his findings extends much further. There are situations in which continuing to play whatever role we are playing at the time—passenger, army private, student—will

handicap us. The knowledge that that is so can make it easier for us to shed our roles when the time is right to do that.

SECTION THREE: SOME RECENT FINDINGS

Recent findings about unconscious influences on behavior are interesting too. In this section, we review evidence of these influences. While a superficial reading of this evidence suggests a pessimistic view about the human capacity for self-control, we offer grounds for optimism.

Melissa Bateson and colleagues (Bateson et al. 2006) conducted an experiment in the office of the Psychology Department at the University of Newcastle. The office keeps coffee, tea, and milk on hand. Department members pay for the drinks by voluntarily depositing money in an “honesty box.” Bateson and colleagues tweaked this system in an interesting way. On a cupboard door located above the honesty box and the drink-making supplies, they posted an instruction sheet with the following suggestions: 30 pence for tea, 50 pence for coffee, 10 pence for milk. In addition to these suggestions, the sheet included an image: either a pair of eyes looking at the observer or flowers. The experiment ran for ten weeks. Each week the experimenters switched the image and recorded the amount of money given that week.

Contributions to the honesty box reliably tracked the change in images. Each time the experimenters replaced the flowers with watching eyes, contributions rose. And each time they replaced the watching eyes with flowers, contributions dropped. On average, department members contributed 2.76 times more money when the eyes were watching (Bateson et al. 2006, p. 412).

Why would the image of eyes have such a powerful effect on behavior? Bateson and colleagues speculate that the eyes “induce a perception in participants of being watched” (2006, p. 413). Importantly, based on information that the human perceptual system is highly sensitive to social stimuli such as eyes, they speculate that this perception is largely non-conscious. The idea is that the presence of the watching eyes activates non-conscious “reputational concerns” that motivate increased contributions (p. 413).

This interpretation receives some support from a study conducted by Mary Rigdon and colleagues. Rigdon et al. 2009 had participants play a version of the “dictator game” in which one participant (the Dictator) was given \$10 and told to indicate on a decision sheet how much—in \$1 increments—he or she wished to give to

another participant (the Recipient). Because it is well-known that social cues about Recipients (e.g., being told the Recipient's surname) influence how much Dictators give, in this study Dictators and Recipients were kept anonymous. The only social cue the Dictators received came in the form of three dots, located in the center of the decision sheet, just above the place where they were to indicate the amount they would give. One group saw three dots arranged to resemble watching eyes and a nose: two dots on top, one the on bottom. A second group saw the three dots arranged in a neutral configuration: one dot on top, two dots on the bottom. This is obviously a very minimal social cue. But dots arranged in the watching eyes configuration are known to activate the part of the brain responsible for face recognition, the fusiform face area (Tong et al. 2000); and Rigdon et al. hypothesized that even this minimal cue would influence the amount given by Dictators.

They were right, but with a twist. Male participants gave significantly more when they saw the watching eyes configuration of dots. On average, male Dictators in the watching eyes condition gave \$3.00; male Dictators in the neutral dots condition gave \$1.41. Furthermore, male Dictators in the watching eyes condition gave \$1 or more 79% of the time, compared with only 37% in the neutral dots condition (Rigdon et al. 2009, p. 362). It seems that for many males, the three dots totally changed their giving behavior.

Female participants' giving patterns were not influenced by the dots at all. Rigdon et al. explain the difference by pointing to other studies using the dictator game that indicate that, on the whole, female Dictators give much more than male Dictators. Since female Dictators "seem to already view the choice problem . . . as a social allocation task" the watching eyes should not stimulate pro-social behavior (2009, p. 363). The opposite is true for male Dictators, however. They tend to use their anonymity to their own economic advantage. Thus, for male Dictators, "Processing the stimulus ultimately activates the fusiform face area of the brain, making the environment seem—at a pre-conscious level, perhaps accessible to the decision-making process but not to introspection—less anonymous and hence less socially distant" (p. 363).

These kinds of studies certainly seem to favor a pessimistic view regarding the human capacity for self-control. Having your fusiform face area stimulated by three dots that vaguely resemble eyes and a nose might activate non-conscious processes that cause you to give money when otherwise you would not have—especially if you were already inclined to be stingy. However, as we noted in section two, it seems

that knowledge about these kinds of non-conscious influences has the potential to enhance our capacity to counteract them. Someone who knows that the presence of eyes, or even the feeling of being watched, tends to cause certain patterns of behavior, might be able to counteract the non-conscious influence. Imagine receiving a flyer in the mail asking for your donation to some politician's campaign, and imagine noticing the image of a face or of eyes in the center of the page. Now that you know that perceiving eyes influences giving behavior, it is possible that different processes will be activated in you. Perhaps you will think more deliberately about how much money you have to donate this month. Perhaps, recognizing an attempt at manipulation, you will toss the flyer into the trash. That is up to you. Our point is that knowledge about non-conscious processes that influence behavior has the potential to mitigate the influence those processes have by activating more explicit processes.

Is there hard evidence that supports our optimism about knowledge? An interesting line of research concerns the behavioral influence of implicit attitudes—attitudes agents possess, but of which they are rarely aware. The most popular measure of such attitudes is the implicit association test (IAT). On a typical IAT—for example, one measuring implicit attitudes towards black and white people—a participant sits in front of a computer screen and is asked to categorize stimuli by pressing one of two keys. There are four types of stimuli—black people's faces, white people's faces, negative words, and positive words—and two response keys. For each response key, researchers pair a black face with either a positive or a negative word and a white face with either a positive or a negative word. This yields two types of response situation: "compatible" (i.e., black face/negative word or white face/positive word) and "incompatible" (i.e., black face/positive word or white face/negative word). Examples of positive words are "joy," "laughter," "love," and "peace"; negative words include "evil," "failure," "nasty," and "terrible." Participants see a series of stimuli and react appropriately: for example, black person's face (press left key), positive word (press left key), negative word (press right key), white person's face (press right key). Researchers then measure how long it takes participants to press the relevant key in different response situations. If they find, as they often do, that it takes longer to categorize a black person's face when the response key maps a black face with a positive word, they conclude that an implicitly negative attitude towards black people exists. As Fiedler and Bluemke explain, "Whoever participated in an IAT, swearing not to be prejudiced at all against Blacks, will have found it nevertheless much easier to use the same re-

sponse for White and positive and for Black and negative than vice versa” (2005, p. 307).

It is commonly assumed that the influence of implicit attitudes on behavior is important and occurs largely by way of non-conscious processing. To take one example, Neil Levy remarks that “implicit attitudes probably explain some incidents involving lethal force. Priming with black faces raises the likelihood that agents will identify ambiguous stimuli or non-gun tools as guns . . . This fact may partially explain why police are more likely to use deadly force when confronted with black suspects” (2012, p. 9).

The assumption that implicit attitudes influence behavior in ways beyond an agent’s conscious control is reflected in assumptions about the IAT. It is assumed that at very short time scales of one second or less, participants cannot consciously control their reactions to stimuli. As Fiedler and Bluemke note, “It is this apparent lack of control or impossibility to counteract the IAT effect that has nourished the claim that an unobtrusive instrument [for measuring implicit attitudes] has been found, which does not lend itself to controlled responding” (2005, p. 307).

The assumption at issue about conscious control is false. Fiedler and Bluemke (2005) gave German participants an IAT that measured negative attitudes against Turks and then asked them to take it again and to try “to avoid a result that would indicate a negative implicit attitude against Turks” (p. 308). Note that even if participants actually have implicitly negative attitudes against Turks, there are, in principle, two ways they might avoid such a result. First, they might slow down their responses in compatible response situations (e.g., Turkish face/negative word). Second, they might speed up their responses in incompatible response situations (e.g., Turkish face/positive word). The slowing down response is one way to exercise control over the influence of one’s implicit attitude: one “beats the test” by pretending that it takes one just as long to categorize stimuli in the compatible as in the incompatible condition. The speeding up strategy is another way to exercise control over the influence of one’s attitude: by making a successful conscious effort, one nullifies the influence of one’s implicit attitude on one’s response time in incompatible response situations.

Fiedler and Bluemke’s participants seemed to utilize both strategies. On their second time through the IAT, they slowed down responses in compatible response situations and sped up responses in incompatible situations (2005, p. 310, Table 2). Fiedler and Bluemke did not predict the latter result, and they were surprised that participants were able to speed up responses (p. 315). They suggested that the speed

up might have been due to practice alone (p. 315). An alternative possibility, not considered by Fiedler and Bluemke, is that participants' familiarity with the test, coupled with an intention to speed up responses, led to the speed up. And there is evidence that bears on it, as we will explain.

Xiaoqing Hu and colleagues (2012) had participants take an IAT and then take it again. On the second trial, they separated participants into four groups. Group 1 simply repeated the IAT to test for the influence of task repetition. Group 2 repeated the incompatible response block of the IAT three times to test for the influence of practice. Group 3 was explicitly instructed to speed up their responses in incompatible response situations. Group 4 was told the same thing as group 3, and they were also given more time to practice; they repeated the incompatible response block three times, just like group 2.

If a conscious intention to speed up responses is to be effective, one would expect group 3 to respond faster than group 1 in the incompatible response conditions. One would also expect group 4 to respond faster than group 2 in the incompatible response conditions. This is what happened (Hu et al. 2012, p. 3, Table 1). Group 3 improved response time by 168 ms (from 902 ms to 734 ms), while group 1 improved response time only by 45 ms (from 950 ms to 905 ms). Compared with group 2, group 4 significantly improved response time as well. Practice certainly seemed to help: group 2 improved response time by 80 ms (from 922 ms to 842 ms). But group 4 improved response time by 215 ms (from 858 ms to 643 ms).

That both a conscious intention and training in speeding up responses had large effects on behavior constitutes important evidence in favor of our optimism. Participants were, in effect, asked to control the influence of implicit attitudes on behavior at a very rapid time scale—less than a second. Participants informed about the influence of implicit attitudes on behavior were able to successfully control the influence of these implicit attitudes. This directly counters the common assumption that implicit attitudes influence behavior in ways not susceptible to conscious control. Knowledge about effects on agents that normally fly under the radar of agents' consciousness can give people the power to weaken those effects. The fact that relevant knowledge can do this at such rapid time scales is striking, and it speaks against a pessimistic perspective on agential control.

SECTION FOUR: IMPLEMENTATION INTENTIONS
AND THE ZOMBIE HYPOTHESIS

John Kihlstrom reports that some social psychologists “embrace and promote the idea that automatic processes dominate human experience, thought, and action to the virtual exclusion of everything else” (2008, p. 168). If he is exaggerating, he is not exaggerating much: for example, Daniel Wegner contends that conscious intentions are never among the causes of corresponding actions (2002, 2004, 2008). However, there is additional good evidence that conscious intentions do important work. If this were a book, we might try to catalogue such evidence. Instead we will concentrate on a particular body of work with clear connections to self-control. We want to continue countering the impression that, as science correspondent Sandra Blakeslee put it in a New York Times article, “in navigating the world and deciding what is rewarding, humans are closer to zombies than sentient beings much of the time” (as quoted in Kihlstrom 2008, p. 163).

Some of Milgram’s descriptions of the excessively obedient behavior he observed are similar to descriptions of akratic action, sometimes defined as un-compelled, intentional action contrary to the agent’s conscious better judgment (see Mele 2012, p. 3). For example, he asserts that “Some subjects were totally convinced of the wrongness of what they were doing” (1974, p. 10) and that “many subjects [who continue shocking] make the intellectual decision that they should not give any more shocks” (p. 148). It is possible that Milgram did not have akratic action in mind. He might have thought that the participants at issue were compelled to act as they did. Milgram completes the first sentence quoted in this paragraph with the words “but *could not* bring themselves to make an open break with authority,” and the second sentence ends with the words “they are frequently *unable* to transform this conviction into action.” (The emphasis is ours in both cases.) But he might have meant that the participants found it very difficult to do what they believed was right and failed to do the right thing.

The flip side of akratic action is enkratic action, action exhibiting self-control in the face of pressure to act contrary to one’s better judgment. A large body of work on “implementation intentions” provides encouragement concerning our prospects for self-control (for reviews, see Gollwitzer 1999 and Gollwitzer and Sheeran 2006) while also countering the idea that conscious intentions have virtually no effect on intentional action.

Implementation intentions, as Peter Gollwitzer describes them, “are subordinate to goal intentions and specify the when, where, and how of responses leading to goal attainment” (1999, p. 494). They “serve the purpose of promoting the attainment of the goal specified in the goal intention.” In forming an implementation intention, “the person commits himself or herself to respond to a certain situation in a certain manner.”

In one study of participants “who had reported strong goal intentions to perform a BSE [breast self-examination] during the next month, 100% did so if they had been induced to form additional implementation intentions” (Gollwitzer 1999, p. 496). In a control group of people who also reported strong goal intentions to do this but were not asked to form implementation intentions, only 53% performed a BSE. Participants in the former group were asked to state in writing “where and when” they would perform a BSE during the next month. These statements expressed implementation intentions.

The featured future task in another study was “vigorous exercise for 20 minutes during the next week” (Gollwitzer 1999, p. 496). “A motivational intervention that focused on increasing self-efficacy to exercise, the perceived severity of and vulnerability to coronary heart disease, and the expectation that exercise will reduce the risk of coronary heart disease raised compliance from 29% to only 39%.” When this intervention was paired with the instruction to form relevant implementation intentions, “the compliance rate rose to 91%.”

In a third study reviewed in Gollwitzer 1999, drug addicts who showed symptoms of withdrawal were divided into two groups. “One group was asked in the morning to form the goal intention to write a short curriculum vitae before 5:00 p.m. and to add implementation intentions that specified when and where they would write it” (p. 496). The other participants were asked “to form the same goal intention but with irrelevant implementation intentions (i.e., they were asked to specify when they would eat lunch and where they would sit).” Once again, the results are striking: although none of the people in the second group completed the task, 80% of the people in the first group completed it.

Many studies of this kind are reviewed in Gollwitzer 1999, and Gollwitzer and Paschal Sheeran report that “findings from 94 independent tests showed that implementation intentions had a positive effect of medium-to-large magnitude . . . on goal attainment” (2006, p. 69). These results provide evidence that the presence of relevant distal implementation intentions significantly increases the probability that agents

will execute associated distal “goal intentions” in a broad range of circumstances. In the experimental studies that Gollwitzer reviews, participants are explicitly asked to form relevant implementation intentions, and the intentions at issue are consciously expressed (1999, p. 501). (It should not be assumed, incidentally, that all members of all of the control groups lack conscious implementation intentions. Indeed, for all anyone knows, many members of the control groups who executed their goal intentions consciously made relevant distal implementation decisions.)

Research on implementation intentions certainly suggests that one useful technique for mastering anticipated motivation not to do what one judges it best to do later—for example, exercise next week or finish writing a C.V. by the end of the day—is simply to decide, shortly after making the judgment, on a very specific plan for so doing. Of course, what works against relatively modest motivational opposition might not work when the opposition is considerably stronger, as it may often be in the case of addicts’ desires for their preferred drugs (see Webb et al. 2009).

We are not suggesting that implementation intentions provide a solution to the problems encountered by participants in the classic studies we have discussed. Our purpose in this section has been to offer some grounds for not being overly impressed by the zombie hypothesis about human beings and some support for our optimism about human prospects for self-control (for additional support, see Mele 2012, ch. 5). The key to dealing with the bystander effect, the power of roles, and excessive obedience, we have suggested, is education. Sometimes, knowledge is power.

SECTION FIVE: OUR STRATEGY

A brief discussion of our strategy in this article is in order. We began by noting that the assumption that people are capable of making and acting on informed, conscious decisions has been challenged by work in the human sciences. The primary challenge that we selected for discussion comes from the situationist literature in psychology. In sections 1 and 2 we reviewed some classic situationist findings and sketched a case for an optimistic perspective on them. Readers will have noticed that our optimism has a cautious tone: for example, we claimed that “once a cause of harmful action or inaction is brought to light, prospects for amelioration may become brighter.” We suggested that a public that is educated about the bystander effect is “less likely to display it,” and we made comparable suggestions about undue or excessive obedience to authority and the effects of agents’ roles.

In section three, we turned to recent findings. We had two aims there. One was to move the discussion into the sphere of current scientific research. The other and more important aim was to call attention to some hard evidence that knowledge about the effects of unconscious processes on behavior can help people to counteract those effects. The work on implicit attitudes that we discussed was useful in this connection, even though it is not situated in the situationist literature. One looks for evidence where one thinks one has a good chance of finding it.

An important part of the assumption that has been our topic in this article is that conscious intentions and decisions can have an effect on behavior, and, more specifically, that they (or their physical correlates) can issue in corresponding actions. The most vigorous opponent of this assumption is Daniel Wegner (2002, 2004, 2008). Our primary aim in section four was to present hard evidence that the opposition is mistaken—evidence from research on implementation intentions. This research is not situated in the situationist tradition; but, again, it makes sense to look for evidence where one believes one is likely to find it.

At this point, the discussion could have taken a metaphysical turn. Is it conscious implementation intentions qua conscious intentions that do the work or is the work instead done by their physical correlates? We opted against discussing that issue here, and we direct readers interested in the topic to Mele 2009, pp. 146-48.

An empirical issue that we did not pursue merits at least a brief mention. There is evidence that pertinent implementation intentions reduce automatic stereotyping (Stewart and Payne 2008). This finding links our discussion of implementation intentions in section four to our discussion of implicit attitudes in section three. (See Fine 2006 for a review of evidence that “automatic social processes can come to be importantly constrained by prior controlled cognitive processes” (p. 85), including prior processes involving implementation intentions (pp. 92-93).) But, again, our primary aim in section four was to counter the idea that conscious intentions play no role in the production of corresponding actions and that automaticity rules here.

We chose not to critique various studies of unconscious influences on behavior. But we would be remiss if we did not note the existence of relevant critiques. For example, in a review of research on unconscious influences on decision making, Ben Newell and David Shanks argue that in various alleged demonstrations of such influence, “inadequate procedures for assessing awareness, failures to consider artifactual explanations of ‘landmark’ results, and a tendency to uncritically accept conclusions that fit with our intuitions have all contributed to unconscious influences being as-

cribed inflated and erroneous explanatory power in theories of decision making” (n.d., p.1).

As we see it, good scientific research on the effects of unconscious processes on behavior should be encouraged, as should good critiques of that work. We value knowledge of the springs of human behavior for its own sake, but such knowledge has instrumental value as well, including the value that the cautiously optimistic perspective we have developed highlights. Sometimes, we said, knowledge is power. Here is another way to put it: sometimes, forewarned is forearmed. It is knowledge about actual broad effects that we have in mind—for example, the bystander effect. We know of no direct evidence that informing people about the bystander effect can influence their behavior in bystander situations. But we do believe that people should be informed about the effect, and we hope that well-presented information will have a positive influence on behavior. In fact, this is part of our reason for describing the bystander study in section one, and we had similar motivation for our presentation of the other classic studies described there. Our understanding is that the readership of this journal will include people who have only a vague familiarity with situationism and classic situationist studies, and the *Journal of Practical Ethics* is a good place for theoretical papers that also pursue some practical aims with ethical significance.

Someday, we may try to catalogue harmful or counterproductive unconscious processes that can be consciously counteracted. We have discussed implicit attitudes in this connection and we have mentioned that implementation intentions have been found to reduce automatic racial stereotyping. Here are two more examples. First, there is evidence that unconscious gender bias in hiring decisions can be counteracted by consciously settling on hiring criteria before the candidates’ gender is disclosed (Uhlmann and Cohen 2005). Second, the confirmation bias—the tendency to search (in memory and the world) more often for confirming than for disconfirming instances of a hypothesis one is testing and to recognize confirming instances more readily—can be counteracted by consciously taking the perspective of someone whose job it is to find violations of a rule (Gigerenzer and Hug 1992). We find research on unconscious processes useful both for what it tells us about how human beings function and for what it might tell us about how human beings can function better. Of course, we are by no means suggesting that most unconscious processes are counterproductive. Some are; many are very useful.

CONCLUSION

The classic situationist studies that we discussed are disconcerting. One response is pessimism about human agency: some may conclude that intentional human action is driven primarily by forces that fly under the radar of consciousness and that we have little insight, as agents, into why we do what we do. Not only have we discussed some evidence to the contrary, but we also have provided grounds for an optimistic view according to which knowledge about situational influences can improve human agents' prospects for dealing rationally with them.

Acknowledgments: We are grateful to two anonymous referees for useful comments on an earlier version of this article. This article was made possible through the support of a grant to Mele from the John Templeton Foundation. The opinions expressed in this article are our own and do not necessarily reflect the views of the John Templeton Foundation.

REFERENCES

- Bateson, M., D. Nettle, and G. Roberts. 2006. Cues of Being Watched Enhance Cooperation in a Real-World Setting. *Biology Letters*, 2: 412-414.
- Darley, J., and B. Latané. 1968. Bystander Intervention in Emergencies: Diffusion of Responsibility. *Journal of Personality and Social Psychology*, 8: 377-383.
- Doris, J. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Fiedler, K. and M. Bluemke. 2005. Faking the IAT: Aided and Unaided Response Control on the Implicit Association Tests. *Basic and Applied Social Psychology*, 27: 307-316.
- Fine, C. 2006. Is the Emotional Dog Wagging its Rational Tail, or Chasing it?. *Philosophical Explorations*, 9: 83-98.
- Gigerenzer, G. and K. Hug. 1992. Domain-Specific Reasoning: Social Contracts, Cheating, and Perspective Change. *Cognition*, 43: 127-171.
- Gollwitzer, P. 1999. Implementation Intentions. *American Psychologist*, 54: 493-503.
- Gollwitzer, P. and P. Sheeran. 2006. Implementation Intentions and Goal Achievement: A Meta-Analysis of Effects and Processes. *Advances in Experimental Social Psychology*, 38: 69-119.
- Haney, C., W. Banks, and P. Zimbardo. 1973. Interpersonal Dynamics of a Simulated Prison. *International Journal of Criminology and Penology*, 1: 69-97.

Harman, G. 1999. Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error. *Proceedings of the Aristotelian Society*, 99: 315-331.

Hu, X., J.P. Rosenfeld, and G.V. Bodenhausen. 2012. Combating Automatic Autobiographical Associations: The Effect of Instruction and Training in Strategically Concealing Information in the Autobiographical Implicit Association test. *Psychological Science*, DOI: 10.1177/0956797612443834, 1-7.

Kamtekar, R. 2004. Situationism and Virtue Ethics on the Content of Our Character. *Ethics*, 114: 458-491.

Kihlstrom, J. 2008. The Automaticity Juggernaut—or Are We Automaton After All?. In: J. Baer, J. Kaufman, and R. Baumeister, eds. *Are We Free? Psychology and Free Will*. New York: Oxford University Press.

Levy, N. 2012. Consciousness, Implicit Attitudes, and Moral Responsibility. *Noûs*, DOI: 10.1111/j.1468-0068.2011.00853.x, 1-22.

Libet, B. 1985. Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. *Behavioral and Brain Sciences*, 8: 529-66.

——— 1999. Do We Have Free Will?. *Journal of Consciousness Studies*, 6: 47-57.

——— 2004. *Mind Time*. Cambridge, Mass.: Harvard University Press.

Lieberman, M. 2005. Principles, Processes, and Puzzles of Social Cognition. *NeuroImage*, 28: 746-756.

Mele, A. 2009. *Effective Intentions*. Oxford: Oxford University Press.

——— 2012. *Backsliding*. Oxford: Oxford University Press.

——— n.d. Free Will and Substance Dualism: The Real Scientific Threat to Free Will?. In: W. Sinnott-Armstrong, ed. *Moral Psychology, Volume 4: Free Will and Responsibility*. Cambridge, Mass: MIT Press.

Milgram, S. 1963. Behavioral Study of Obedience. *The Journal of Abnormal and Social Psychology*, 67: 371-78.

——— 1965. Some Conditions of Obedience and Disobedience to Authority. *Human Relations*, 18: 57-76.

——— 1974. *Obedience to Authority*. New York: Harper & Row.

Rigdon, M., K. Ishii, M. Watabe, and S. Kitayama. 2009. Minimal Social Cues in the Dictator Game. *Journal of Economic Psychology*, 30: 358-367.

Newell, B. and D. Shanks. n.d. Unconscious Influences on Decision Making: A Critical Review. *Behavioral and Brain Sciences*.

Sabini, J. and M. Silver. 2005. Lack of Character? Situationism Critiqued. *Ethics*, 115: 535-562.

Sreenivasan, G. 2002. Errors about Errors: Virtue Theory and Trait Attribution. *Mind*, 111: 47-68.

Stewart, B. and K. Payne. 2008. Bringing Automatic Stereotyping Under Control: Implementation Intentions as Efficient Means of Thought Control. *Personality and Social Psychology Bulletin*, 34: 1332-1345.

Tong, F., K. Nakayama, M. Moscovitch, O. Wienrib, and N. Kanwisher. 2000. Response Properties of the Human Fusiform Face Area. *Cognitive Neuropsychology*, 17: 257-279.

Uhlmann, E. and G. Cohen. 2005. Constructed Criteria: Redefining Merit to Justify Discrimination. *Psychological Science*, 16: 474-480.

Webb, T., P. Sheeran, and A. Luszczynska. 2009. Planning to Break Unwanted Habits: Habit Strength Moderates Implementation Effects on Behavior Change. *British Journal of Social Psychology*, 48: 507-523.

Wegner, D. 2002. *The Illusion of Conscious Will*. Cambridge, Mass. MIT Press.

——— 2004. Précis of 'The Illusion of Conscious Will'. *Behavioral and Brain Sciences*, 27: 649-659.

——— 2008. Self is Magic. In: J. Baer, J. Kaufman, and R. Baumeister, eds. *Are We Free? Psychology and Free Will*. New York: Oxford University Press.

Zimbardo, P. n.d. "Stanford Prison Experiment": <http://www.prisonexp.org/>.

Zimbardo, P., C. Haney, W. Banks, and D. Jaffe. 1973. The Mind Is a Formidable Jailer: A Pirandellian Prison. *The New York Times Magazine*, Section 6 (April 8: 38-60). <http://www.prisonexp.org/>.