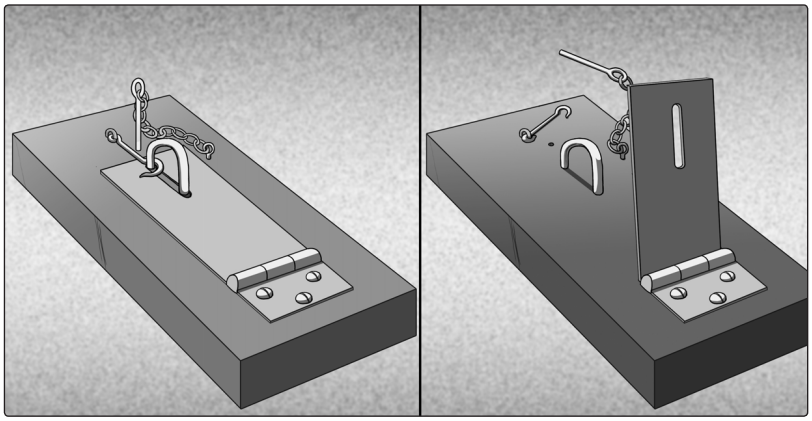
INTRODUCTION

The Puzzling Puzzles of Harry Harlow and Edward Deci

In the middle of the last century, two young scientists conducted experiments that should have changed the world—but did not.

Harry F. Harlow was a professor of psychology at the University of Wisconsin who, in the 1940s, established one of the world's first laboratories for studying primate behavior. One day in 1949, Harlow and two colleagues gathered eight rhesus monkeys for a two-week experiment on learning. The researchers devised a simple mechanical puzzle like the one pictured below. Solving it required three steps: (a) pull out the vertical pin, (b) undo the hook, and (c) lift the hinged cover. Pretty easy for you and me, far more challenging for a thirteen-pound lab monkey.



Harlow's puzzle in the starting (left) and solved (right) positions. Illustrations by Rob Ten Pas.

The experimenters placed the puzzles into the monkeys' cages to observe how they reacted—and to prepare them for tests of their problem-solving prowess at the end of the two weeks. But almost immediately, something strange happened. Unbidden by any outside urging and unprompted by the experimenters, the monkeys began playing with the puzzles with focus, determination, and what looked like enjoyment. And in short order, they began figuring out how the contraptions worked. By the time Harlow tested the monkeys on days 13 and 14 of the experiment, the primates had become quite adept. They solved the puzzles frequently and quickly; two-thirds of the time they cracked the code in less than sixty seconds.

Now, this was a bit odd. Nobody had taught the monkeys how to remove the pin, slide the hook, and open the cover. Nobody had rewarded them with food, affection, or even quiet applause when they succeeded. And that ran counter to the accepted notions of how primates—including the bigger-brained, less hairy primates known as human beings—behaved.

Scientists then knew that two main drives powered behavior. The

first was the biological drive. Humans and other animals ate to sate their hunger, drank to quench their thirst, and reproduced to satisfy their carnal urges. But that wasn't happening here. "Solution did not lead to food, water, or sex gratification," Harlow reported.[1]

But the only other known drive also failed to explain the monkeys' peculiar behavior. If biological motivations came from within, this second drive came from without—the rewards and punishments the environment delivered for behaving in certain ways. This was certainly true for humans, who responded exquisitely to such external forces. If you promised to raise our pay, we'd work harder. If you held out the prospect of getting an A on test, we'd study longer. If you threatened to dock us for showing up late or for incorrectly completing a form, we'd arrive on time and tick every box. But that didn't account for the monkeys' actions either. As Harlow wrote, and you can almost hear him scratching his head, "The behavior obtained in this investigation poses some interesting questions for motivation theory, since significant learning was attained and efficient performance maintained without resort to special or extrinsic incentives."

What else could it be?

To answer the question, Harlow offered a novel theory—what amounted to a *third* drive: "The performance of the task," he said, "provided intrinsic reward." The monkeys solved the puzzles simply because they found it gratifying to solve puzzles. They enjoyed it. The joy of the task was its own reward.

If this notion was radical, what happened next only deepened the confusion and controversy. Perhaps this newly discovered drive—Harlow eventually called it "intrinsic motivation"—was real. But surely it was subordinate to the other two drives. If the monkeys were rewarded—with raisins!—for solving the puzzles, they'd no doubt perform even better. Yet when Harlow tested that approach, the monkeys actually made *more* errors and solved the puzzles *less*

frequently. "Introduction of food in the present experiment," Harlow wrote, "served to disrupt performance, a phenomenon not reported in the literature."

Now, this was *really* odd. In scientific terms, it was akin to rolling a steel ball down an inclined plane to measure its velocity—only to watch the ball float into the air instead. It suggested that our understanding of the gravitational pulls on our behavior was inadequate—that what we thought were fixed laws had plenty of loopholes. Harlow emphasized the "strength and persistence" of the monkeys' drive to complete the puzzles. Then he noted:

> It would appear that this drive . . . may be as basic and strong as the [other] drives. Furthermore, there is some reason to believe that [it] can be as efficient in facilitating learning.[2]

At the time, however, the prevailing two drives held a tight grip on scientific thinking. So Harlow sounded the alarm. He urged scientists to "close down large sections of our theoretical junkyard" and offer fresher, more accurate accounts of human behavior.[3] He warned that our explanation of why we did what we did was incomplete. He said that to truly understand the human condition, we had to take account of this third drive.

Then he pretty much dropped the whole idea.

Rather than battle the establishment and begin offering a more complete view of motivation, Harlow abandoned this contentious line of research and later became famous for studies on the science of affection.[4] His notion of this third drive bounced around the psychological literature, but it remained on the periphery—of behavioral science and of our understanding of ourselves. It would be two decades before another scientist picked up the thread that Harlow had so provocatively left on that Wisconsin laboratory table.