

DOI:10.1145/1897852.1897872

**Effective countermeasures depend on first understanding how users naturally fall victim to fraudsters.**

**BY FRANK STAJANO AND PAUL WILSON**

# Understanding Scam Victims: Seven Principles for Systems Security

FROM A HOLISTIC security engineering point of view, real-world systems are often vulnerable to attack despite being protected by elaborate technical safeguards. The weakest point in any security-strengthened system is usually its human element; an attack is possible because the designers thought only about their strategy for responding to threats, without anticipating how real users would react.

We need to understand how users behave and what traits of that behavior make them vulnerable, then design systems security around them. To gain this

knowledge, we examine a variety of scams, distilling some general principles of human behavior that explain why the scams work; we then show how they also apply to broader attacks on computer systems insofar as they involve humans. Awareness of the aspects of human psychology exploited by con artists helps not only the public avoid these particular scams but also security engineers build more robust systems.

Over nine series of the BBC TV documentary *The Real Hustle* (<http://www.bbc.co.uk/realhustle/>) Paul Wilson and Alexis Conran researched the scams most commonly carried out in Britain and, with Jessica-Jane Clement, replicated hundreds of them on unsuspecting victims while filming the action with hidden cameras. The victims were later debriefed, given their money back, and asked for their consent to publish the footage so others would learn not to fall for the same scams (see the sidebar “Representative Scams” to which we refer throughout the main text.)

The objective of the TV show was to help viewers avoid being ripped off by similar scams. Can security researchers do more? By carefully dissecting dozens of scams, we extracted seven recurring behavioral patterns and related principles exhibited by victims and exploited by hustlers. They are not merely small-scale opportunistic scams (known as “short cons”) but in-

## » key insights

- **We observed and documented hundreds of frauds, but almost all of them can be reduced to a handful of general principles that explain what victims fall for.**
- **These principles cause vulnerabilities in computer systems but were exploited by fraudsters for centuries before computers were invented and are rooted in human nature.**
- **Users fall prey to these principles not because they are gullible but because they are human. Instead of blaming users, understand that these inherent vulnerabilities exist, then make your system robust despite them.**



herent security vulnerabilities of the human element in any complex system. The security engineer must understand them thoroughly and consider their implications toward computer and system security.

### **Distraction Principle**

*While we are distracted by what grabs our interest, hustlers can do anything to us and we won't notice.*

The young lady who falls prey to the recruitment scam is so engrossed in her job-finding task that she totally fails to even suspect that the whole agency might be a fraud.

Distraction is at the heart of innumerable fraud scenarios. It is also a fundamental ingredient of most magic performances,<sup>5</sup> which is not surprising if we see such performances as a “benign fraud” for entertainment purposes. Distraction is used in all cases

involving sleight of hand, including pickpocketing and the special “throw” found in the Monte.

The very presence of “sexy swindler” Jess among the hustlers owes to Distraction, as well as to Need and Greed (discussed later), since sex is such a fundamental human drive. The 2000 computer worm “ILOVEYOU,” which reportedly caused \$5 billion–\$8 billion damage worldwide, exploited these two principles.

In computing, the well-known tension between security and usability is also related to Distraction. Users care only about what they want to access and are essentially blind to the fact that “the annoying security gobbledygook” is there to protect them. Smart crooks exploit this mismatch to their advantage; a lock that is inconvenient to use is often left open.

Distraction also plays a role in the

“419,” or Nigerian, scam. The hustler, posing as a Nigerian government officer with access to tens of millions of dollars of dodgy money, wants the mark to help transfer the money out of the country in exchange for a slice of it. When the mark accepts the deal, the hustler demands some amount of advance money to cover expenses. New unexpected expenses come up repeatedly, always with the promise that the money is just about to be transferred. These “convincers” keep the mark focused solely on the huge sum he is promised to receive.

Are only unsophisticated 419 victims gullible? Abagnale<sup>1</sup> showed the Distraction principle works equally well on highly educated CTOs and CIOs. In 1999, he visited a company full of programmers frantically fixing code to avert the Y2K bug. He asked the executives how they found all the program-

mers and was told “these guys from India” knew computers well and were inexpensive. But, Abagnale thought, any dishonest programmer from an offshore firm fixing Y2K problems could also easily implant a backdoor...

People focused on what they want to do are distracted from the task of protecting themselves. Security engineers who don’t understand this principle have already lost the battle.

### Social Compliance Principle

*Society trains people to not question authority. Hustlers exploit this “suspension of suspiciousness” to make us do what they want.*

The jeweler in a jewelry-shop scam gratefully hands over necklace and cash when “policeman” Alex says they’re needed as evidence, believing him saying they’ll be returned later.

Access control to sensitive databases may involve an exploitable human element. For example, social-engineering-expert Mitnick<sup>7</sup> impersonates a policeman to nothing less than a law-enforcement agency. He builds up credibility and trust by exhibiting knowledge of the lingo, procedures, and phone numbers. He makes the clerk consult the National Crime Information Center database and acquires confidential information about a chosen victim. His insightful observation is that the police and military, far from being a tougher target, are inherently more vulnerable to social engineering as a consequence of their strongly ingrained respect for rank.

Social Compliance is the foundation for phishing. For example our banks, which hold all our money, order us to type our password, and, naturally, we do. It’s difficult to fault nontechnical users on this one if they fail to notice the site was only a lookalike. Note the conflict between a bank’s security department telling customers “never click on email links” and the marketing department of the same bank sending them clickable email advertisements for new financial products, putting the customers in double jeopardy.

System architects must coherently align incentives and liabilities with overall system goals. If users are expected to perform sanity checks rather than blindly follow orders, then social protocols must allow “challenging the authority”; if, on the contrary, users are expected to obey authority unquestioningly, those with authority must relieve them of liability if they obey a fraudster. The fight against phishing and all other forms of social engineering can never be won unless this principle is understood.

### Herd Principle

*Even suspicious marks let their guard down when everyone around them appears to share the same risks. Safety in numbers? Not if they’re all conspiring against us.*

In the Monte, most participants are skills. The whole game is set up to give the mark confidence and make him think: “Yes, the game looks dodgy, but other people are winning money,” and

“Yes, the game looks difficult, but I did guess where the winning disc was, even if that guy lost.” Skills are a key ingredient.

In online auctions, a variety of frauds are possible if bidders are in cahoots with the auctioneer. EBay pioneered a reputation system in which bidders and auctioneers rate each other through public feedback. But fraudsters might boost their reputations through successful transactions with skills. Basic reputation systems are largely ineffective against skills.

In online communities and social networks, multiple aliases created by certain participants to give the impression that others share their opinions are indicated as “sock-puppets.” In political elections, introducing fake identities to simulate grass-roots support for a candidate is called “astroturfing.” In reputation systems in peer-to-peer networks, as opposed to reputation systems in human communities, multiple entities controlled by the same attacker are called “Sybils.” The variety of terms created for different contexts testifies to the wide applicability of the Herd principle to many kinds of multi-user systems.

### Dishonesty Principle

*Our own inner larceny is what hooks us initially. Thereafter, anything illegal we do will be used against us by fraudsters.*

In the Monte, the skills encourage the mark to cheat the operator and even help him do it. Then, having fleeced the mark, the operator pretends to notice the mark’s attempt at cheating, using it as a reason for closing the game without giving him a chance to argue.

When hustlers sell stolen goods, the implied message is “It’s illegal; that’s why you’re getting such a good deal,” so marks won’t go to the police once they discover they’ve been had. The Dishonesty Principle is at the core of the 419; once a mark realizes it’s a scam, calling the police is scary because the mark’s part of the deal (essentially money laundering) was in itself illegal and punishable. Several victims have gone bankrupt, and some have even committed suicide, seeing no way out of this tunnel.

The security engineer must be aware of the Dishonesty Principle. A

Principles to which victims respond, as identified by three sets of researchers.

| Principle                                   | Cialdini (1985–2009) | Lea et al. (2009) | Stajano-Wilson (2009) |
|---------------------------------------------|----------------------|-------------------|-----------------------|
| Distraction                                 |                      | ~                 | •                     |
| Social Compliance (a.k.a. “Authority”)      | •                    | ○                 | ○                     |
| Herd (a.k.a. “Social Proof”)                | •                    |                   | ○                     |
| Dishonesty                                  |                      |                   | •                     |
| Kindness                                    | ~                    |                   | •                     |
| Need and Greed (a.k.a. “Visceral Triggers”) | ~                    | •                 | ○                     |
| Scarcity (related to our “Time”)            | •                    | ○                 | ~                     |
| Commitment and Consistency                  | •                    | ○                 |                       |
| Reciprocation                               | •                    |                   | ~                     |

- First identified this principle
- Also lists this principle
- ~ Lists a related principle

number of attacks on the system go unreported because the victims won't confess to their "evil" part in the process. When corporate users fall prey to a Trojan horse program purporting to offer, say, free access to porn, they have strong incentives not to cooperate with the forensic investigations of system administrators to avoid the associated stigma, even if the incident affected the security of the whole corporate network. Executives for whom righteousness is not as important as the security of their enterprise might consider reflecting such priorities in the corporate security policy, perhaps by guaranteeing discretion and immunity from "internal prosecution" for victims who cooperate with forensic investigations.

### Kindness Principle

*People are fundamentally nice and willing to help. Hustlers shamelessly take advantage of it.*

This principle is, in some sense, the dual of the Dishonesty Principle, as perfectly demonstrated by the Good Samaritan scam. In it, marks are hustled primarily because they volunteer to help. It is loosely related to Cialdini's Reciprocation Principle (people return favors)<sup>2</sup> but applies even in the absence of a "first move" from the hustler. A variety of scams that propagate through email or social networks involve tear-jerking personal stories or follow disaster news (tsunami, earthquake, hurricane), taking advantage of the generous but naïve recipients following their spontaneous kindness before suspecting anything. Many "social engineering" penetrations of computer systems<sup>7</sup> also rely on victims' innate helpfulness.

### Need and Greed Principle

*Our needs and desires make us vulnerable. Once hustlers know what we want, they can easily manipulate us.*

Loewenstein<sup>4</sup> speaks of "visceral factors such as the cravings associated with drug addiction, drive states (such as hunger, thirst, and sexual desire), moods and emotions, and physical pain." We say "Need and Greed" to refer to this spectrum of human needs and desires—all the stuff we really want, regardless of moral judgement. In the 419 scam, what matters most is not necessarily the mark's greed but

his or her personal situation; if the mark is on the verge of bankruptcy, needs major surgery, or is otherwise in dire straits, then questioning the offer of a solution is very difficult. In such cases the mark is not greedy, just depressed and hopeful. If someone prays every day for an answer, an email message from a Nigerian Prince might seem like the heaven-sent solution.

The inclusion of sexual appetite as a fundamental human need justifies, through this principle, the presence of a "sexy swindler" in most scams enacted by "the trio." As noted, the Need and Greed Principle and the Distraction Principle are often connected; victims are distracted by (and toward) that which they desire. This drive is exploited by a vast proportion of fraudulent email messages (such as those involving length enhancers, dates with attractive prospects, viruses, and Trojans, including ILOVEYOU).

An enlightened system administrator once unofficially provided a few gigabytes of soft porn on an intranet server in order to make it unnecessary for local users to go looking for such material on dodgy sites outside the corporate firewall, thereby reducing at the same time connection charges and exposure to malware.

If we want to con someone, all we need to know is what they want, even if it doesn't exist. If security engineers do not understand what users want, and that they want it so badly they'll go to any lengths to get it, then they won't understand what drives users and won't be able to predict their behavior. Engineers always lose against fraudsters who do understand how they can lead their marks. This brings us back to the security/usability trade-off: Lecturing users about disabling ActiveX or Flash or Javascript from untrusted sites is pointless if these software components are required to access what users want or need (such as their online social network site or online banking site or online tax return site). Fraudsters must merely promise some enticing content to enroll users as unwitting accomplices who unlock the doors from inside.

The defense strategy should also include user education; as the *Real Hustle* TV show often says, "If it sounds too good to be true, it probably is."

### Time Principle

*When under time pressure to make an important choice, we use a different decision strategy, and hustlers steer us toward one involving less reasoning.*

In the ring-reward rip-off, the mark is made to believe he must act quickly or lose the opportunity. When caught in such a trap, it's very difficult for people to stop and assess the situation properly.

Unlike the theory of rational choice, that is, that humans take their decision after seeking the optimal solution based on all the available information, Simon<sup>8</sup> suggested that "organisms adapt well enough to 'satisfice'; they do not, in general, 'optimize'."

They may "satisfice," or reach a "good-enough" solution, through simplifying heuristics rather than the complex, reasoned strategies needed for finding the best solution, despite heuristics occasionally failing, as studied by Tversky and Kahneman.<sup>10</sup>

Though hustlers may have never formally studied the psychology of decision making, they intuitively understand the shift. They know that, when forced to take a decision quickly, a mark will not think clearly, acting on impulse according to predictable patterns. So they make their marks an offer they can't refuse, making it clear to them that it's their only chance to accept it. This pattern is evident in the 419 scam and in phishing ("You'll lose access to your bank account if you don't confirm your credentials immediately") but also in various email offers and limited-time discounts in the gray area between acceptable marketing techniques and outright swindle. As modern computerized marketing relies more and more on profiling individual consumers to figure out how to press their buttons, we might periodically have to revise our opinions about which sales methods, while not yet illegal, are ethically acceptable.

From a systems point of view, the Time Principle is particularly important, highlighting that, due to the human element, the system's response to the same stimulus may be radically different depending on the urgency with which it is requested. In military contexts this is taken into account by wrapping dangerous situations that require rapid response (such as challeng-

# Representative Scams

Since 2006, the *Real Hustle* TV show has recreated hundreds of scams during which Paul, Alex, and Jess defrauded unsuspecting victims before hidden cameras. Here are five instructive ones:

In the lingo of this peculiar “trade,” the victim of the scam is the mark, the perpetrator is the operator, and any accomplice pretending to be a regular customer is a shill.

**Monte.** This classic scam involves an operator manipulating three cards (or disks or shells: there are many variations), one of which wins, while the other two lose. The operator shows the player the cards, turns them over face down, then moves them around on the table in full view. Players must follow the moves and put money on the card they believe to be the winner. The operator pays out an equal amount if the player guessed correctly or otherwise pockets the player’s money.

Technically, at the core of the scam is a sleight-of-hand trick whereby the

operator undetectably switches two cards. One might therefore imagine the basic scam to consist of performing a few “demo runs” where marks are allowed to guess correctly, then have them bet with real money and at that point send the winning card elsewhere.

But this so-called “game” is really a cleverly structured piece of street theater designed to attract passersby and hook them into the action. The sleight-of-hand element is actually least important; it is the way marks are manipulated, rather than the props, that brings in the money. It’s all about the crowd of onlookers and players (all shills) betting in a frenzy and irresistibly sucking marks into wanting a piece of the action.

The Monte is an excellent example that nothing is what it seems, even if the marks think they know what to expect. Many people claim to be able to beat the game, purely because they understand the mechanics of the secret move. But it’s impossible to tell whether an experienced

operator has made the switch. More important, even if the cards were marked in some way, there is absolutely no way for a legitimate player to secure a win; should a mark consistently bet on the correct position, then other players, actually shills, would over-bet him, “forcing” the operator to take the larger bet. This frustrates the mark, who often increases his bet to avoid being topped. One shill will then pretend to help the mark by bending a corner of the winning card while the operator is distracted, making the mark think he has an unbeatable advantage. This is a very strong play; marks have been seen to drop thousands of dollars only to find the bent card is actually a loser. While mixing the cards, it is possible for a skilled operator to switch the cards and switch the bend from one card to another.

The idea that one can beat the game at all reveals a key misunderstanding—that, in fact, it is not a game in the first place. Monte mobs never pay out to the



From right to left: Paul, with Alex as a shill, scams two marks at the three-shells game (one of several variants of the Monte).



From right to left: Paul and Alex haggle with the mark over the reward in the Ring Reward Rip-off.



Alex, flashing a fake police badge, pretends to arrest Jess in the Jewelry Shop Scam.

ALL IMAGES COURTESY OF OBJECTIVE PRODUCTIONS

ing strangers at a checkpoint or being ordered to launch a nuclear missile) in special “human protocols” meant to enforce, even under time pressure, some of the step-by-step rational checks the heuristic strategy would otherwise omit.

The security architect must identify the situations in which the humans in the system may suddenly be put under time pressure by an attacker and whether the resulting switch in decision strategy might open a vulnerability. This directive applies to anything from retail situations to stock trading and online auctions and from admitting visitors into buildings to handling medical emergencies. Devising a human protocol to guide and pace the response of the potential victim toward the desired goal may be an adequate safeguard and also relieve the victim from stressful responsibility.

## Related Work

While a few narrative accounts of scams and frauds are available, from Maurer’s study of the criminal world<sup>6</sup> that inspired the 1973 movie *The Sting* to the autobiographical works of notable fraudsters,<sup>1,7</sup> the literature contains little about systematic studies of fraudsters’ psychological techniques. But we found two notable exceptions: Cialdini’s outstanding book *Influence: Science and Practice*,<sup>2</sup> based on undercover field research, revealed how salespeople’s “weapons of influence” are remarkably similar to those of fraudsters; indeed, all of his principles apply to our scenario and vice versa. Meanwhile, Lea et al.<sup>3</sup> examined postal scams, based on a wealth of experimental data, including interviews with victims and lexical analysis of fraudulent letters. Even though our approaches were quite different, our

findings are in substantial agreement. The table here summarizes and compares the principles identified in each of these works.

## Conclusion

We supported our thesis—that systems involving people can be made secure only if designers understand and acknowledge the inherent vulnerabilities of the “human factor”—with three main contributions:

First is a vast body of original research on scams, initially put together by Wilson and Conran. It started as a TV show, not as a controlled scientific experiment, but our representative write-up<sup>9</sup> still offers valuable firsthand data not otherwise available in the literature;

Second, from these hundreds of scams, we abstracted seven principles. The particular principles are not that important, and others have found

marks; they keep all the money moving between the shills and the operator. The marks are allowed to place a bet only if it's already a loser. Having studied Monte all over the world, we can say it's nothing short of a polite way to mug people.

**Ring reward rip-off.** The gorgeous Jess buys a cheap ring from a market stall for \$5. She then goes to a pub and seductively befriends the barman (the mark). She makes it obvious she's very rich; showing off to her friend (a shill), she makes sure the mark overhears that she just received this amazing \$3,500 diamond ring for her birthday. She then leaves.

Paul and Alex arrive at the pub, posing as two blokes having a pint. Jess then phones the pub, very worried, calling her friend the barman by name, saying she lost that very precious ring. Could he check if it's there somewhere? The mark checks, and, luckily, a customer (Paul) found the ring. However, instead of handing it over, Paul demands

a reward. The barman gets back to the phone and Jess, very relieved to hear the ring is there, says, without prompting, that she'll give \$200 to the person who found it. But the barman goes back to Paul and says the reward is only \$20. That's when the hustlers know they've got him; he's trying to make some profit for himself. Paul haggles a bit and eventually returns the ring to the barman for \$50. The mark is all too happy to advance the money to Paul, expecting to get much more from Jess. Jess, of course, never calls back.

A convicted criminal proudly says he once made a \$2,000 profit with this particular hustle.

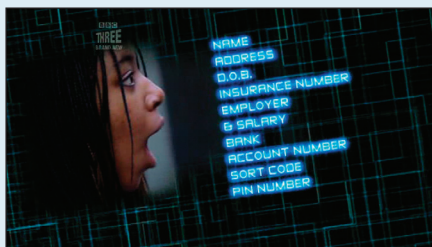
**Jewelry-shop scam.** Jess attempts to buy an expensive necklace but is "arrested" by Alex and Paul posing as plainclothes police officers who expose her as a well-known fraudster, notorious for paying with counterfeit cash. The "cops" collect as evidence the "counterfeit" (actually genuine) cash

and, crucially, the necklace, which will, of course, "be returned." The jeweler is extremely grateful the cops saved her from the evil fraudster.

Ironically, as Jess is taken away in handcuffs, the upset jeweler spits out a venomous "Bitch! You could have cost me my job. You know that?"

**Recruitment scam.** Hustlers set up a fake recruitment agency and, as part of the sign-on procedure, collect all of the applicants' personal details, including mother's maiden name, date of birth, bank-account details, passport number, even PIN—by asking them to protect their data with a four-digit code, as many people memorize only one PIN and use it for everything. With this loot, the hustlers are free to engage in identity theft on everyone who came in for an interview.

**Good Samaritan scam.** In a parking lot, Jess has jacked up her car but seems stuck. When another car stops nearby, she politely asks the newcomers to help her change the tire, which they do. Apologizing for her cheekiness, she then also asks them if they could get into their car, as she's been out in the cold for a while and is freezing. The gentleman gives her the keys to his car (required to turn on the heat) and, while the marks are busy changing her tire, she drives off with the car. But didn't Jess just lose her original car? No, because it wasn't hers to start with; she just jacked up a random one in the parking lot. To add insult to injury, the marks will also have some explaining to do when the real owners of the car arrive.



A mark, debriefed by accompanying TV crew, is dismayed to learn the hustlers just got hold of all her sensitive personal details in the Recruitment Scam.



From right to left: Jess gets two marks to change her tire before tricking them into handing over their own car keys in the Good Samaritan Scam.

slightly different ones. What matters is recognizing the existence of a small set of behavioral patterns that ordinary people exhibit and that hustlers have been exploiting forever; and

Third, perhaps most significant, we applied the principles to a more general systems point of view. The behavioral patterns are not just opportunities for small-scale hustles but also vulnerabilities of the human component of any complex system.

Our message for the system-security architect is that it is naïve to lay blame on users and whine, "The system I designed would be secure, if only users were less gullible." The wise security designer seeking a robust solution will acknowledge the existence of these vulnerabilities as an unavoidable consequence of human nature and actively build safeguards that prevent their exploitation.

## Acknowledgments

Special thanks to Alex Conran for co-writing the TV series and to Alex and Jess Clement for co-starring in it. Thanks to Joe Bonneau, danah boyd, Omar Choudary, Saar Drimer, Jeff Hancock, David Livingstone Smith, Ford-Long Wong, Ross Anderson, Stuart Wray, and especially Roberto Viviani for useful comments on previous drafts. This article is updated and abridged from the 2009 technical report<sup>9</sup> by the same authors. **C**

## References

1. Abagnale, F.W. *The Art of the Steal: How to Protect Yourself and Your Business from Fraud*. Broadway Books, New York, 2001.
2. Cialdini, R.B. *Influence: Science and Practice, Fifth Edition*. Pearson, Boston, MA, 2009; (First Edition 1985).
3. Lea et al. *The Psychology of Scams: Provoking and Committing Errors of Judgement*. Technical Report OFT1070. University of Exeter School of Psychology. Office of Fair Trading, London, U.K., May 2009.
4. Loewenstein, G. Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision*

*Processes* 65, 3 (Mar. 1996), 272–292.

5. Macknik, S.L., King, M., Randi, J., Robbins, A., Teller, Thompson, J., and Martinez-Conde, S. Attention and awareness in stage magic: Turning tricks into research. *Nature Reviews Neuroscience* 9, 11 (Nov. 2008), 871–879.
6. Maurer, D.W. *The Big Con: The Story of the Confidence Man*. Bobbs-Merrill, New York, 1940.
7. Mitnick, K.D. *The Art of Deception: Controlling the Human Element of Security*. John Wiley & Sons, Inc., New York, 2002.
8. Simon, H.A. Rational choice and the structure of the environment. *Psychological Review* 63, 2 (Mar. 1956), 129–138.
9. Stajano, F. and Wilson, P. *Understanding Scam Victims: Seven Principles for Systems Security*. Technical Report UCAM-CL-TR-754. University of Cambridge Computer Laboratory, Cambridge, U.K., 2009.
10. Tversky, A. and Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science* 185, 4157 (Sept. 1974), 1124–1131.

**Frank Stajano** (frank.stajano@cl.cam.ac.uk) is a university senior lecturer in the Computer Laboratory of the University of Cambridge, Cambridge, U.K.

**Paul Wilson** (info@conartist.tv) is an expert on cheating, award-winning conjuror, and magic inventor. He works in film and television in London and Los Angeles.

© 2011 ACM 0001-0782/11/0300 \$10.00