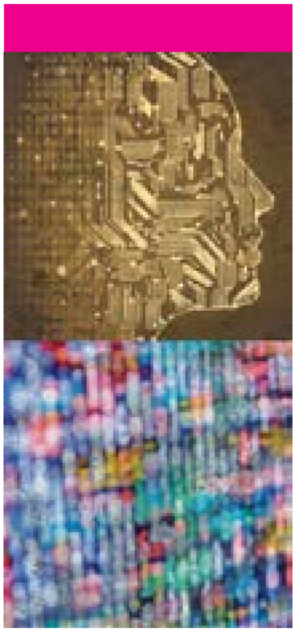
McKinsey&Company

McKinsey on Payments

Special Edition on Advanced Analytics in Banking

Volume 11 Number 28 August 2018

“All in the mind”: Harnessing psychology and analytics to counter bias and reduce risk



The management of risk in financial services is about to be transformed. A recent McKinsey paper identified six structural trends that will reshape the function in the next decade. Five are familiar—they concern regulation, costs, customer expectations, analytics, and digitization—but one is less so: debiasing. That means using insights from psychology and behavioral economics, combined with advanced analytical methods, to take the bias out of risk decisions. The institutions pioneering this approach have seen tremendous benefits: for instance, banks adopting psychological interventions in consumer collections have achieved a 20 to 30 percent increase in the amount collected.[1]

Tobias Baer
Vijay D’Silva

The interest in debiasing is growing as psychological research uncovers more and more subconscious effects that influence our decision making. Meanwhile, an explosion in data availability is providing businesses with an abundant flow of information for their analytic engines. Not all the data theoretically available can be exploited, for legal and privacy as well as technical reasons. But institutions still have a massive amount of underused data that they can mine, using an increasingly sophisticated array of advanced analytics techniques, to develop behavioral segmentations and predictive models. With these foundations in place, they can go on to design powerful interventions to tackle bias.

Take the example of a bank using a recursive neural network to extract customer profiles from credit-card transaction data. One profile that emerges is of a cardholder who clocks up dozens of low-value transactions at a convenience store every week. The customer’s habit of making multiple repeat visits at odd hours—seemingly for only one or two items at a time—suggests a lack of forward planning. Seen through a psychometric lens, the customer seems to be exhibiting poor impulse control and a lack of conscientiousness, traits that are likely to determine which types of decision bias this customer can be expected to manifest.

Compare this profile with that of a cardholder who completes one big supermarket transaction at more or less the same time every Friday evening, with little or no evidence of convenience-store shopping in between. That profile is indicative of a well-organized person who plans ahead. It’s likely that the first customer would benefit from financial products designed to help customers who struggle to meet their financial obligations— such as a credit card with weekly rather than monthly payment installments—whereas the second customer would probably have no need of them. And if, say, the bank is considering ways to motivate cardholders to pay off delinquent credit, its knowledge that customers with the first cardholder’s profile are likely to prioritize immediate consumption over clearing their debts will help it design suitable incentives to counter this tendency.

Analytics-driven psychological insights like these can be a spur to tremendous value creation. This article considers some of the most

[1] For a comprehensive discussion of the psychological levers that can be used to improve performance in consumer debt collection, see Tobias Baer, “Behavioral insights and innovative treatments in collections,” *McKinsey on Risk*, Number 5, March 2018.

common biases in business decision making and looks in detail at three areas where debiasing can reap rich rewards: credit underwriting, consumer debt collection, and asset management.

Uncovering biases in business

Biases are predispositions of a psychological, sociological, or even physiological nature that can influence our decision making (see sidebar, "A quick guide to common biases"). They often operate subconsciously, outside the logical processes that we like to believe govern our decisions. They are frequently regarded as flaws, but this is both wrong and unfortunate. It's wrong because biases are an inevitable side-effect of the mechanics our brains need to achieve their astonishing speed and efficiency in making tens of thousands of decisions a day. And it's unfortunate because the negative perception of biases leads us to believe we are immune to them—a bias in itself, known as overconfidence, exhibited by the 93 percent of US drivers who believe themselves to be among the nation's top 50 percent.

Even if we accept that biases may influence our decisions, we might assume that successful organizations have developed processes to keep them in check. But experience indicates otherwise. For example, academic research has found that ego depletion materially affects the work of judges, doctors, and crime investigators, and our own research has revealed how it affects credit officers' decisions, manifesting itself in tangible business metrics such as credit approval rates. When financial institutions work to counter bias in judgmental underwriting—in small business credit, for example—they can typically cut credit losses by at least 25 percent, and even

A quick guide to common biases

Heuristic biases are computational shortcuts taken by the brain to achieve lightning-fast, almost effortless decision making. Thanks to the Nobel Prize–winning work of Daniel Kahneman and Richard Thaler, these biases have become more widely understood in recent years. More than a hundred have been identified, ranging from the relatively familiar loss aversion to the less well-known Hawthorne effect. For practical business purposes, five groups of biases are key:

- **Action-oriented biases** prompt us to act with less forethought than is logically necessary or prudent. They include *excessive optimism* about outcomes and the tendency to underestimate the likelihood of negative results; *overconfidence* in our own or our group's ability to affect the future; and competitor neglect, the tendency to disregard or underestimate the response of competitors.

- **Interest biases** arise where incentives within an organization or project come into conflict. They include misaligned individual incentives, unwarranted emotional attachments to elements of the business (such as legacy products), and differing perceptions of corporate goals, such as how much weight to assign to particular objectives.

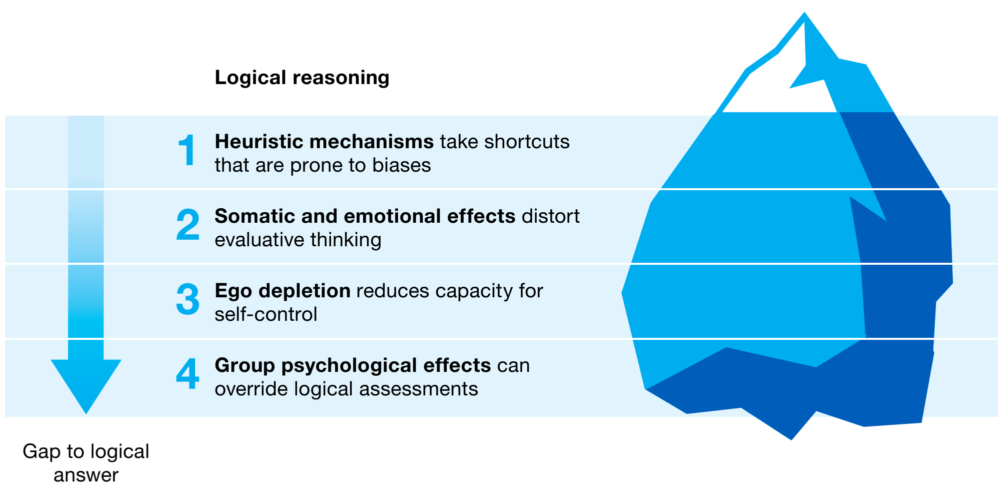
- **Pattern-recognition biases** cause us to see nonexistent patterns in information. They include *confirmation bias*, in which we overvalue evidence that supports a favored belief and discount evidence to the contrary; availability bias, in which we misperceive likelihoods of events because we recall one type of event much more easily (and hence frequently) than others; management by example, in which we rely unduly on our own experiences when making decisions; and *false analogies*, faulty thinking based on incorrect perceptions and the treatment of dissimilar things as similar.

- **Stability biases** predispose us toward inertia in an uncertain environment. They include *anchoring without sufficient adjustment*, in which we tie actions to an initial value but fail to adjust when new information becomes available; loss aversion, the fear that makes us more risk-averse than logic would dictate; the *sunk-cost fallacy*, where our future course of action is influenced by the unrecoverable costs of the past; and *status-quo bias*, the preference for keeping things as they are when there is no immediate pressure to change.
- **Social biases** arise from our preference for harmony over conflict, or even constructive challenge. They include *groupthink*, in which the desire for consensus prevents us making a realistic appraisal of alternative courses of action, and *sunflower management*, the tendency for group members to fall into line with their leaders' views.

For all their importance, however, heuristic biases represent only the tip of the iceberg as far as subconscious influences on our decisions are concerned. Exhibit A illustrates other factors that lie deep below the surface. Somatic and emotional effects tinker with the parameterization of our brain, and can be triggered by factors as diverse as blood-sugar level, smells, or mood: for instance, if our blood sugar is low, we (quite reasonably) estimate that completing a given task, such as climbing a mountain, will take us longer. Ego depletion, a form of mental fatigue, leads us to move from logical thinking to unconscious short-cuts that favor easy default decisions. And group psychological effects override rational decision making out of a deep-seated fear of ostracism.

Exhibit A

Psychologists and neuroscientists have discovered many forces that cause decisions to gravitate away from logical considerations.



Source: McKinsey analysis

as much as 57 percent in one case.

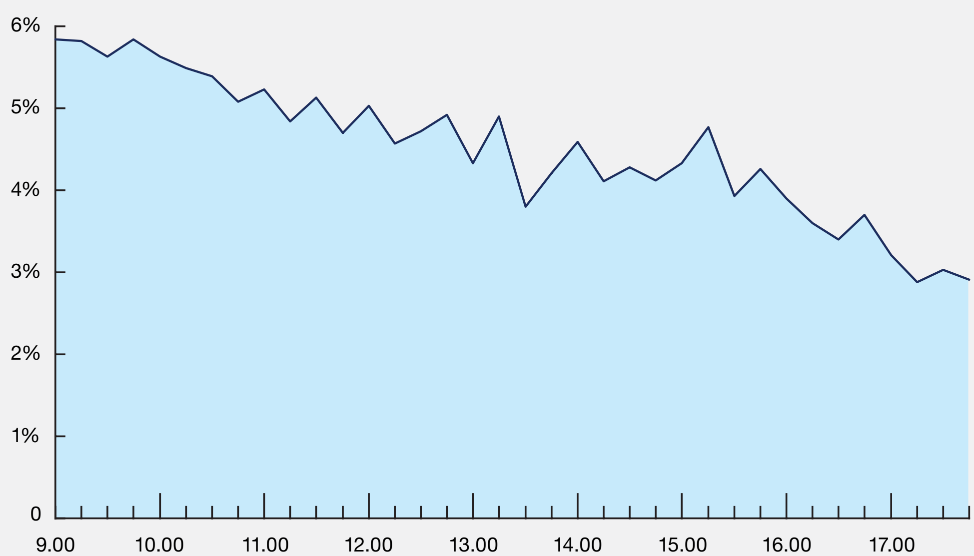
For lenders, an area particularly ripe for debiasing is debt collections, where biases can shape the behavior of collectors and customers alike. Consider how collectors handle calls with recalcitrant customers. Over the course of a call, they need to make numerous split-second decisions that expose them to the full gamut of biases, such as anchoring and over-optimism, as well as somatic effects and ego depletion. Whether they persist in trying to elicit a promise to pay or give up and move on to the next delinquent account may partly depend on the time of day. The effectiveness of collectors' calls dwindles over the course of the working day as ego depletion sets in (Exhibit 1). The good news is that companies aware of this phenomenon can make adjustments in collectors' working environment to help counter it.

And when it comes to customers with overdue accounts, leading financial institutions are harnessing a plethora of psychological insights to encourage payment. This often means making targeted interventions that increase customers' motivation to pay, help those with low self-control to keep their commitments, and respect individuals' need for agency (and thereby avoid triggering what psychologists call "reactance"). A credit-card provider could, for instance, present high-risk customers with a late-fee waiver or a gift card from a favorite shop that they would lose if they didn't make a payment. Framing the

Exhibit 1

Call effectiveness dwindles over the course of the day through ego depletion as collectors tire.

Case example; percent of calls eliciting promise to pay, by time of day



Source: McKinsey analysis

How can financial institutions tackle biases?

The questions organizations need to consider include:

The decision type: High or low frequency, formal or informal

Formal high-frequency decisions, such as credit underwriting or standard manual fraud checks, lend themselves to analytical solutions coupled with "industrial strength" psychological interventions. For example, a bank that usually asks about the frequency of CFO changes in the past three years—a question that may be susceptible to the availability bias—could instead design a simple table prompting credit officers to construct a timeline for pertinent data points.

Formal low-frequency decisions—such as approvals of new lending products or a credit committee's quarterly recalibration of the PD rating model that drives underwriting and risk-based pricing—call for decision processes to be redesigned to support logical thinking and ensure adequate challenge. Analytical modeling is often helpful here. One US bank used four different econometric models to produce four distinct default-rate forecasts in an elegant effort to counteract groupthink and introduce automated "devil's advocates" into the discussion.

Informal decisions, such as a supervisor's override on a policy violation, may first have to be formalized before any intervention can be deployed. A review of historical losses may shed light on a few decision types that warrant such an investment, such as debt collectors' decisions to give up on difficult accounts. If a bank wants a collector to spend longer than usual on a call to a particular customer, for instance, it could flag up an above-average incentive payment in a pop-up on the collector's screen.

Who to target and how

Institutions need to use behavioral segmentation to distinguish which groups are affected by which primary biases, and which personality traits determine the choice of countermeasure. In consumer debt collection, for instance, the psychological need for agency can cause customers to resist resolution if they feel they have been put on the spot by a call from an assertive collector. An invitation to restructure the debt on a self-service website could effectively overcome this bias. However, this same approach could be disastrous if used to deal with a customer who is biased toward avoidance.

The role of automation

Carefully designed algorithms can not only speed up decisions and take out costs, but also remove biases from a growing range of decision types. But financial institutions must beware of a major trap: building past biases into the algorithm.

offer as a loss for a payment missed, rather than a reward for a payment made, enlists the help of the loss aversion bias and can double the effectiveness of the offer.

Before deciding where and how to use behavioral levers, financial institutions need to consider a range of factors (see sidebar, "How can financial institutions tackle biases?", page 72).

To give a sense of what can be achieved when these techniques are applied in practice, let's now examine what leading institutions have been doing to take bias out of credit underwriting and consumer debt collection. And looking beyond lending, the sidebar "Debiasing asset management" (page 76) describes how firms in an adjacent industry uncovered bias in their investment decisions.

Commercial credit underwriting

Most credit officers possess a strong professional ethic and have honed their skills over years, if not decades. Yet evidence indicates they are just as susceptible as anyone else to decision bias.

One bank with poor performance in its commercial credit underwriting made a retrospective assessment of the predictive value of its judgmental credit ratings using Gini coefficient measures on a scale from 0 (no predictive power) to 100 (perfect prediction). The analysis examined 20 dimensions stipulated by the bank's credit policy, such as management quality and account conduct, and compared judgments made by credit officers with actual defaults observed over the following 12 months. One dimension (account conduct) stood out with a relatively high Gini of 45, but most dimensions had much lower scores (Exhibit 2). By way of comparison, comprehensive best-practice models for rating small businesses can achieve a Gini of 60–75.

In fact, half of the dimensions in the bank's rating model achieved a Gini score of 7 or lower—little better than a roll of the dice—yet the bank was paying them just as much attention as it gave to dimensions with genuine predictive power. For instance, despite scoring a Gini of just 1 in back-testing, shareholder composition was usually discussed in depth in credit memos, and relationship managers were even prompted to ask customers follow-up questions about it. Factoring in such irrelevant dimensions anchored credit officers' overall rating in randomness, dragging it down to a Gini of just 22.

In order to debias its commercial underwriting, the bank had to separate the wheat from the chaff—a systematic process combining analytics with psychological insights. First, the bank replaced fuzzy concepts with carefully chosen sets of proxies for which more objective assessments could be developed. Eliminating factors that were irrelevant, or impossible to assess without crippling bias, would substantially improve the overall credit rating. Second, explicit psychological "guard rails," such as the use of tables to prompt credit officers to plot data along a timeline rather than relying on a customer's spontaneous recall of events, were put in place to safeguard qualitative assessment processes from biases.

Finally, the bank used statistical techniques to validate each redesigned factor and calibrate its weight. As is common in commercial credit portfolios, the bank ran up against the problem of a small sample size. This was compounded by the need to compile additional data manually for the sample used in developing the new assessment, which comprised just 30 to 50 defaulters and the same number of performing debtors. Although such constraints ruled out the statistical techniques most commonly used in credit scoring, such as logistic regression, the bank was able to deploy powerful statistical concepts from social science instead, such as Cohen's d and t-test.

The bank has now been using its qualitative credit rating, with minimal annual adjustments, for more than a decade, scoring an overall Gini between 60 and 80 every year, even during the financial crisis.

Consumer collections

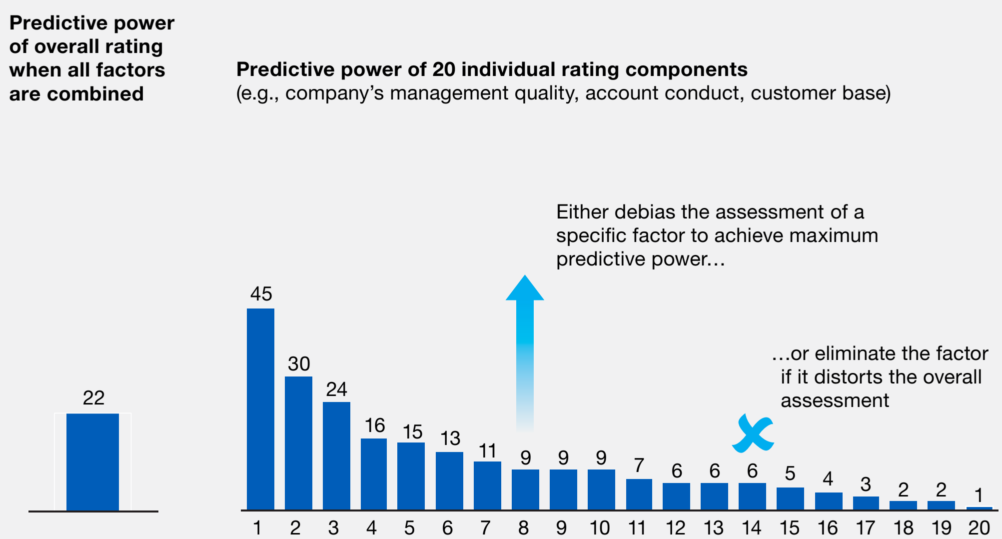
A recent McKinsey survey of 420 US consumers with credit delinquencies sheds light on some of the decision biases that contribute to non-payment. For instance, many consumers are unable to resist the temptation of immediate consumption—an example of what's known as "hyperbolic discounting"—and so they struggle to manage money through a monthly cycle. A third of those surveyed expressed a preference for a schedule that would allow payment every week or every other week, either because it would fit better with their paydays or because smaller, more frequent payments would be less painful and easier to manage than monthly bills.

Understanding how consumers decide what to pay and when is particularly important

Exhibit 2

One bank's credit-rating model contained factors that were subject to bias or had little or no predictive power.

Case example; predictive power of judgmental ratings assigned by credit officers, measured by Gini coefficient: 0 = useless (random), 100 = perfect prediction



To debias credit memos, institutions can replace lengthy prose with concise questions, multiple-choice options, and simple tables—which will also streamline assessment, cut costs, and speed up turnaround

Source: McKinsey analysis

when they owe money to more than one lender. Only a third of survey respondents prioritized payments rationally by, say, tackling debts with the highest interest rate first, or seeking to retain their most useful credit card. The remaining two-thirds followed less rational patterns: some apportioned payments equally, others showed loyalty to a particular bank, and yet others paid off the smallest balance first (Exhibit 3). Banks that are aware of such motivations can either reinforce them with tailored payment plans or help customers adjust their rationales—for instance, by breaking down large balances into smaller chunks or milestones.

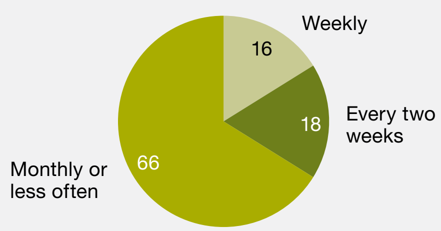
Some leading banks are putting behavioral targeting into practice by applying psychometrics: the factual scoring of a customer's personality profile according to a framework such as the widely used OCEAN Big Five. Such a profile allows banks to micro-target marketing messages not only in origination—choosing the visuals, tag line, and highlighted features to use in a product pitch—but also in debt collection. When applying such an approach, banks often find it helpful to break

Exhibit 3

Research into consumers with credit delinquencies yielded valuable behavioral insights.

Survey of 420 US consumers who have been at least one month overdue

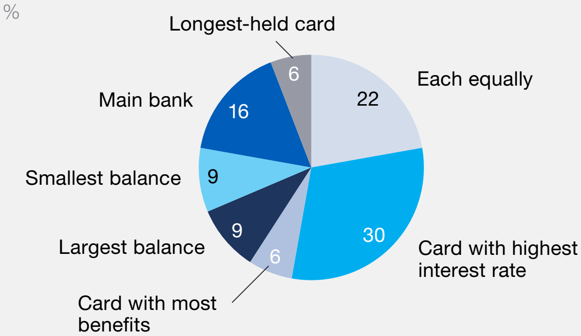
What payment frequency do you prefer?
%



Comments and insights

"I always prefer to pay smaller payments more frequently because it takes the sting out of making a payment. Making a large payment always feels like a punch."

When several accounts are overdue, which do you pay first?[1]
%



Comments and insights

20% of respondents said they have **withheld a planned payment** because of an upsetting call from a collector

38% of respondents had a very positive experience with at least one collector who was **empathetic and genuinely helpful**

[1] Figures do not sum to 100% because of rounding

Source: McKinsey survey

Debiasing asset management

Few industries have subjected their investment decision-making processes to more scrutiny than asset management, yet biases still affect many high-value decisions throughout the lifecycle of individual funds. In the early stages of structuring a new fund's strategy and processes, for instance, stability biases can influence whether an index or some other means is chosen for assessing performance. Interest biases, such as misaligned incentives, need to be monitored to ensure that the long-term interests of unit holders and asset owners are taken into account when funds are managed and promoted.

Leading asset management organizations are becoming increasingly alert to the impact of decision-making biases on fund performance too. A few have adopted an innovative approach to diagnosing bias and its drivers. Working with analytics experts and behavioral scientists, they have applied machine-learning algorithms to their own historical data and discovered clusters of suboptimal investment decisions. Having examined these decisions more closely, they have detected signs of consistent bias in the processes by which the decisions were reached.

When one such organization analyzed its trades, processes, and associated emotions for signs of bias, it found that more than 35 percent of fund managers' decisions were influenced by biases such as loss aversion, anchoring, and what's known as the "endowment effect," in which we attach more value to items that we own. Dan Ariely, a behavioral economist and the best-selling author of Predictably Irrational, notes that this effect kicks in when individuals fall in love with what they already have and focus on what they may lose rather than what they may gain. Such a sentiment can drive fund managers to hold on to stocks for too long and ignore better investment opportunities elsewhere—a trap into which many seasoned investors have fallen.

In one of the funds that this organization examined, the endowment effect had led one fund manager to hold on to 20 percent of positions for too long. The stocks affected had underperformed the relevant index by an average of 25 percent in the 12 months prior to exit. The fund manager acknowledged that he had paid insufficient attention to these stocks, had not rated them as performing badly enough in absolute terms to divest, and could have tried harder to identify better investment opportunities. He admitted that if he had asked himself from time to time whether he would still buy the stocks today, he would have been unlikely to hold on to them for so long. In this case, the value left on the table as a result of the endowment effect was equivalent to 250 to 300 basis points per year.

And this fund manager is not alone. According to Cabot research, institutional investors lose an average of 100 basis points in performance a year as a result of the endowment effect—or 250 basis points in the case of the 10 percent of most-affected funds.

A typical debiasing process is a learning exercise for an asset management fund. By exposing patterns of bias with the help of analytics and then selecting and applying debiasing methods in its investment decisions, the fund will be able to target the specific biases and situations that affect its own investment decisions. From the many interventions available to address every type of bias, it will need to select and customize measures that suit its fund mandate, investment philosophy, team process, culture, and individual personalities.[1]

Magdalena Smith is an expert in McKinsey's London office.

[1] For more on this topic, see Nick Hoffman, Martin Huber and Magdalena Smith, "An analytics approach to debiasing asset-management decisions," McKinsey & Company, December 2017.

down a collections episode into four distinct "moments":

1. *Opening*. When the phone rings, customers must decide whether or not to engage with the bank or card provider. If they pick up, they then have to decide whether to take a defensive or evasive stance or to collaborate in problem solving (for example, by disclosing financial difficulties).

2. *Commitment*. Once collaboration has been established, the collector needs to move the customer toward a promise to pay.

3. *Negotiation*. A major part of the conversation will be a negotiation over the customer's financial limitations and the payment to which he or she is willing to commit.

4. *Follow-through*. Finally, the customer needs to keep the promise to pay—a complex decision with ample opportunities for derailment.

At each of these moments, the customer must decide whether or not to cooperate with the lender, and the lender must try to understand the customer's behavior and identify opportunities to increase the likelihood of repayment, using psychological interventions carefully calibrated to each customer's profile.

In the opening moment, a collector who puts a customer in the right mood (or "positive affect") will increase that person's receptiveness to exploring solutions and self-confidence in resolving the situation. Conversely, creating the opposite mood—negative affect—will impede resolution. One approach that institutions have found effective is to use collectors with profiles similar to those of customers, matching regional dialect, gender, and age. Similarly, requesting a call back via email, text message, or app alert instead of calling the customer directly shows respect for an individual's need for agency. Customers too ashamed or anxious to speak on the phone can sometimes be steered to self-service channels through advertisements on social media.

By telling a customer that the solution being offered has been popular with other clients, collectors can trigger the "herd effect"— one of several techniques proven to move a customer towards a commitment. Anchoring negotiations in a full repayment within a short time-frame will help a customer commit to making the biggest payment they can manage. This not only maximizes recovery for the bank but also protects the customer from unnecessary interest charges and bankruptcy that could result from falling victim to hyperbolic discounting.

Ensuring that customers keep their promise to pay is arguably the hardest part of collections. Again, behavioral segmentation sheds light on the intricate factors determining the decision to follow through—or not—on a promised payment. One justification customers frequently use to rationalize broken promises is the hassle (actual or perceived) involved in making a payment. A quarter of respondents in our survey of US consumers with credit delinquencies said that making payments was a hassle, and a third of this group said they would be more likely to pay if more convenient payment methods were available (Exhibit 4).

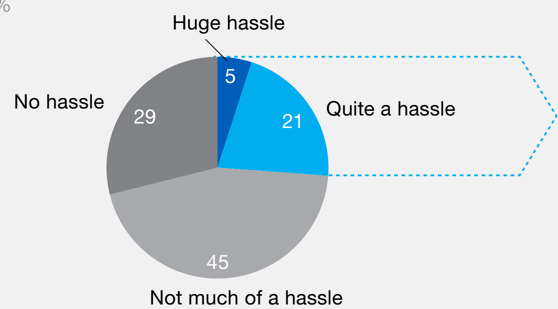
One bank that piloted innovative treatments saw multiple benefits: a 30 percent increase in collections, a 20 percent reduction in write-offs on delinquent debt, a 33 percent

Exhibit 4

For some consumers, payment can be a hassle.

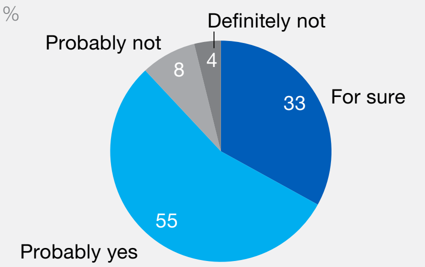
How difficult is it to execute a payment?

%



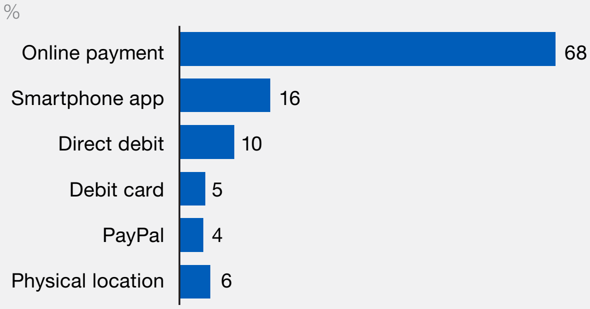
With more convenient payment methods, would you be more likely to make payments?

%



What payment method is easiest for you?

%



Comments

"An option to use a prepaid card or something like that would help. Nine times out of ten if the money gets put in the bank account, it will be taken out by another bill."

"I wish there was an easier way to send payments from my debit account. I hate finding out all the account numbers."

Source: McKinsey survey

fall in delinquencies remaining for late-stage collection, and a 20 percent reduction in the number of customers subsequently relapsing into default (Exhibit 5, page 79). First, the bank used K-means clustering to create an initial segmentation of five behavioral clusters. Next, it used a range of tools including closed-file reviews, psychometric surveys, and interviews to compile an ethnographic profile for each cluster. Finally, it drew on the growing body of psychological research and real-life experience with nudges and other psychological interventions in other industries to design effective treatments..

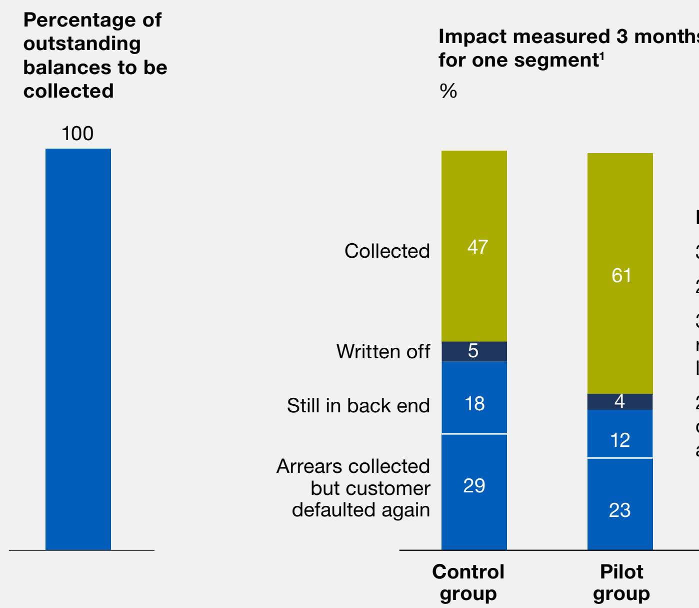
* * *

Being aware of bias and taking deliberate steps to counter it has already proved effective in areas such as gender bias in hiring. A few pioneering financial institutions have adopted a similar approach to debiasing their business decisions and have seen impressive results, such as a 25 to 35 percent reduction in credit losses from improved underwriting and collections. Yet for most institutions, the big prizes have yet to be captured.

The secret lies in combining psychological insights with advanced statistical methods to

Exhibit 5

Tailored treatments based on behavioral segmentation can deliver multiple benefits.



Benefits of behavioral pilot

30% rise in total collected

20% fewer write-offs

33% fewer delinquencies remaining in back end for late-stage collection

20% fewer relapses into default among collected accounts

Behavioral-based prescriptive treatment for this segment

High call priority; no delay in efforts because of messages being left

Thorough inquiry with detailed questions about the customer's situation

Assertive script with no inappropriate "customer service" mindset

Questions about how, where, and when customer will pay help form an "implementation intention" that makes them more likely to keep their promise

[1] Figures may not sum to 100% because of rounding
Source: McKinsey analysis

develop a pragmatic but powerful behavioral segmentation linked to targeted treatments. By introducing creative workarounds into their existing infrastructure, especially in IT implementation, providers can have a new approach up and running in as little as three months, with dramatic effects. Given the impact that early efforts have achieved, it can be only a matter of time before such innovative treatments become the norm.

Tobias Baer conducts psychological research at the University of Cambridge; he is a former partner at McKinsey and a member of our Behavioral Insights Group. **Vijay D'Silva** is a senior partner in McKinsey's New York office.