palgrave pivot

An Introduction to Cognitive Economics

The Science of Mistakes

Andrew Caplin

OPEN ACCESS

palgrave macmillan

An Introduction to Cognitive Economics

Andrew Caplin

An Introduction to Cognitive Economics

The Science of Mistakes

palgrave
macmillan

Andrew Caplin
Department of Economics
New York University
New York, NY, USA

ISBN 978-3-031-73041-2 ISBN 978-3-031-73042-9 (eBook)
https://doi.org/10.1007/978-3-031-73042-9

This work was supported by the Alfred P. Sloan Foundation and the NOMIS Foundation.

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

I thank my collaborators Daniel Martin and Soren Leth-Peterson for their guidance on how to structure the book. I thank my editor Bronwyn Geyer, as well as Danny Goroff, Valerie Whitten, and Ruth Wyatt, for invaluable feedback on early drafts. I thank the Alfred P. Sloan and NOMIS Foundations for their early support of cognitive economics, and the Alfred P. Sloan Foundation for its prescient role in supporting translational research on Cognitive Economics at Work. *I dedicate the book with the greatest love to Ruth, Ann, and Ollie for their loving support and understanding throughout.*

Contents

List of Figures

CHAPTER 1

Introduction to Cognitive Economics

Would you like a child of yours to become a social scientist? On the positive side, we spend our working lives looking for new solutions to the most important economic, political, and social challenges of our time. This book reports on a thirty year project identifying solution methods in the common ground between economics, my mother discipline, and psychology, my field of fascination. On the negative side, office politics is absurd, publishing research can take up to five years, and translation to practice far longer. Good as our ideas might be, it can take many decades for them to make their way into practice. On the whole, there have been better uses of time for those looking to make a difference.

My goal is to convince you that change is coming. A wonderful career awaits those who in the future become captivated by social and economic questions. Foundations are in place for takeoff of *cognitive economics*, the hybrid discipline a small group of dedicated researchers and I have been developing. With this, we will transition to a world in which social scientific advances have a positive impact on the world and earn innovators their just rewards. This book invites your participation in bringing this change about in the shortest possible time. The invitation is open to all. Together we can ensure that even those well-advanced in other careers will be able to join the social scientific enterprise, to the benefit of us all.

Why hurry along the (almost) inevitable? Because the world has not been standing in wait for cognitive economics to come of age. The economy is at the start of the most dramatic technological upheaval since

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_1

the industrial revolution. We are in the early days of the cognitive revolution as we transition to an economy in which all manner of decisions will be made with the help of algorithmic advisors. Workers will create value through their cognitive skills in decision-making, such as creativity in problem-solving and awareness of comparative advantage, rather than through traditional physical or manual labor. Cognitive economics can help us navigate these choppy waters if we act now to accelerate the research process. In this opening chapter I introduce the field, outline its broad applicability, explain the challenges of advancing research, lay out next steps, and explain how readers can help speed progress.

1.1 What Is Cognitive Economics?

Economics isn't solely about money; it's about decision-making. In cognitive economics, we study decisions that may be *mistaken*. Cognitive economics is in effect a *science of mistakes*. That this science is rooted in economics is due to the central role that three model constructs play in making the field scientifically precise.

1. *Utility functions* help us understand what constitutes a mistake by allowing us to compare the decision that was made with one that could have been made and would have been better for the decision-maker. *Better* here means scoring higher on some scale, which is exactly what utility functions define.
2. *Imperfect information* is central because our knowledge about the world is incredibly limited. Do I really need to justify this claim? Humans know next to nothing about almost everything. The vast gap between what we know and what we need to know to make optimal decisions helps explain many of the mistakes we make. The basic theory of decision-making in the face of such information constraints is foundational to the cognitive economic approach to identifying mistakes. The standard model of imperfect information in economics is a *Blackwell experiment* that induces updating of *subjective beliefs* from an ill-informed prior to better informed posterior beliefs.
3. *Costs of learning* are central because, if learning had no costs, the frequency of mistakes would dramatically decrease. We wouldn't need artificial intelligence, teaching, or even basic computing tools if

we could effortlessly process all available information. But learning *is* costly, and these costs are a crucial factor in why we often make sub-optimal decisions.

Pulling together these constructs, cognitive economics places the *production of good decisions* at its core. Interactions between the costs and benefits of cognitive effort play key roles in almost every aspect of our lives. They underlie the services we derive from technologies like smartphones and the efforts we make to plan for the uncertain contingencies we face in all aspects of decision-making. You may find exceptions to this rule of effortful decision-making in the complex environments that we face every day. But these will be exceptions that prove the rule.

1.2 Why Does Cognitive Economics Matter?

At first sight, the study of mistakes might seem removed from bread and butter issues of economics. Nothing could be further from the truth. No area of social scientific research can healthily ignore cognitive constraints. They are universal. By way of illustration, the first part of the book deals with applications of cognitive economics in traditional economic areas related to wealth, choice, and earnings from work.

- Chapter 2 introduces *cognitive household finance* organized around the *life-cycle model* of *wealth accumulation*. It covers *exponential* and *hyperbolic discounting*, *temptation*, and *self-control problems*. It distinguishes the *present bias* of many younger individuals from the *future bias* of many older ones. It discusses *financial literacy*, *financial education*, *financial advisers*, and the power of *financial planning*. It discusses the impact of *cognitive decline* on wealth and *financial abuse of the elderly*. It stresses how difficult it is to provide for future *long-term care* needs and to guard assets against cognitive decline. It indicates the potential value of *shared equity housing finance* for those who wish to receive care in the home.
- Chapter 3 introduces cognitive economic methods for *measuring the quality of communication* and reducing *decision-making mistakes* using *cognitive nudges*. It outlines a case study that shows how adding an index to complex case files improved the *quality of justice* in *Mexican labor arbitration courts* in cases of unfair dismissal.

It introduces *experimental economic methods* to measure the clarity of communication based on *Bloom's educational taxonomy*. It outlines an ongoing application to ensure *privacy disclosures* are as well understood as possible. More broadly it indicates the need for cognitively informed *mandated disclosure regulations* to make *informed consent* more realistic.

- Chapter 4 introduces *cognitive labor economics*. Inspired by the industrial revolution, *standard production functions* model the transformation of physical factors of production into physical outputs. With the growing importance of *higher level decision-making skills*, we need to model inputs as cognitive and outputs as improvements in decision quality. The chapter discusses how to amend *human capital theory* to capture the open-ended tasks that are increasingly in demand. The chapter also introduces recent research measuring and modeling *economic decision-making skills* based on the ability to apply the principle of *comparative advantage*. It introduces evidence that these skills help explain earnings from employment and outlines ongoing research to track these effects in the *Danish population registries*. It also discusses research on the importance of *skills in teamwork* for earnings from work.

There is far more to cognitive economics than shedding light on traditional areas of research. It opens up whole new areas for investigation. What will make cognitive economics such an important science going forward is that it addresses previously unasked questions concerning the integration of AI into our everyday lives. Included in this is how to adapt our methods of teaching in light of the new requirements of work and daily life that the cognitive revolution creates. The second part of the book deals with these vital new areas of social scientific research.

- Chapter 5 introduces *cognitive capital* and *human-AI interactions*. The upcoming cognitive revolution, like its industrial precursor, is defined by a radical change in technology. The machines of today introduce intangible cognitive capital that substitutes for human mental effort. The chapter lays out a three stage *human-AI decision-making pipeline*. This starts with the human ground truth *labeling* of cases, continues with minimization of the *AI loss function*, and involves a final stage of *human-AI collaboration*. The chapter introduces and outlines a field study in which methods of measuring *elicited beliefs* imported from *experimental economics* improve labeling for *cancer detection*. It introduces recent findings showing that standard engineering methods of defining loss functions can fail on their own terms. As a result, there is a *bilateral alignment problem* that requires not only that the AI is trained in human preferences, but that humans understand the AI's costs and constraints.

- Chapter 6 introduces research on how the rise of AI will impact wages as the cognitive revolution unfolds. It sketches recent research suggesting that AI helps those of *lower skill* to close the productivity and earnings gap with those of *higher skill*. It introduces new results suggesting that to so benefit requires those of lower skill to have beliefs that are *well-calibrated* to reality. It outlines the importance of long-run career skills, such as the ability to *search effectively* for new jobs. It also outlines research showing that one can impact students' *choice of education and career* by transmitting more accurate information.
- Chapter 7 introduces applications of cognitive economics to *teaching* and *testing*. It stresses the value of teaching individuals to be *better calibrated*. It outlines the importance of improved *grading methods* for helping students to work more *effectively with AI assistance*. It outlines research findings showing the importance of *better informing students* about the value of their educational options in terms of future employment and earnings. It proposes development of *courses in cognitive economics* to raise the unfortunately low current level of *public discourse* on social scientific questions.

A word of caution. Many people demand answers to the big questions of our time, and a thriving market exists to supply them. However, I will not join that game. This is a *how to*, not a *what is* book. Social science, as it stands, cannot definitively answer any of the big questions. Instead, we are honing wonderful methods for exploring them. That said, I will discuss policy aspects of cognitive economics in Chapter 9 of the book, which makes the case that cognitive economics is ready for prime time. I open by sketching some obvious next research steps. The three topic areas I outline are of particular importance in the transition to a cognitive economy.

- Topic area 1 picks up on the research outlined in Chapter 5. It concerns the design of human-AI decision-making pipelines in important use cases, such as medical decisions, hiring decisions, etc. The goal is to avoid a *cascade of biases* that might otherwise emerge from the Rube Goldberg kluge that currently defines most human-AI decision-making protocols.
- Topic area 2 picks up on the research outlined in Chapter 6. It concerns how well workers and students understand the opportunities and threats associated with the cognitive transition, and how to help them make better educational and career decisions.
- Topic area 3 picks up on the massively under-developed cognitive economic approach to teaching in Chapter 7. It concerns how to understand and teach the skills that will allow people to work effectively with AI. Included is a discussion of how to test students who will increasingly turn to generative AI tools to provide answers.

Chapter 9 closes by highlighting the many policy and business paths that cognitive economic research points to. It offers many ways for policymakers and businesses to tailor cognitive economic research to their own uses. I think this is the just the tip of the iceberg. Ultimately business and policy clusters will form around leading cognitive economic research groups, much as in many of the natural sciences and in data science. Irrelevance is not a virtue.

1.3 An Accelerator for Cognitive Economics

Cognitive economics has deep roots in both economics and cognitive psychology. Chapter 8 presents a history framed in terms of Walt Rostow's *stages of economic growth*. It highlights the important early contribution of Paul Samuelson, who in 1939 introduced *revealed preference theory*, as well as that of Henry Block and Jacob Marschak, who, back in 1959, introduced economists to the psychological idea that choice is impacted by cognitive constraints. It sketches more recent research on *rational inattention*, *bounded rationality*, *efficient coding*, *salience-based attention*, *imperfect memory*, *resource rationality*, and *cognitive hierarchies*. It outlines the ongoing convergence of interest between psychologists and economists on the importance of cognitive constraints.

On the one hand, Chapter 8 tells a story of success as consensus has grown on the importance of cognitive economics. Yet it is also a story of missed opportunity. Utility functions, imperfect information, and costs of learning have been in research focus for decades, and arguably go back to the ancients. The field might easily have taken off in 1959.

I am confident that, in the fullness of time, cognitive economics will develop as the social scientific complement to the cognitive economy. While speed might not matter from a high enough vantage point, from mine it does. On the personal level, I would like to be around when cognitive economics becomes more fully recognized. On the social level, cognitive economics can help us answer urgent questions about the transition to a cognitive economy and catalyze important positive innovations in the world of practice. If we can get ahead of the upcoming changes in economic and social structures, we may build a better future with each of us developing our human potential in ever more diverse and personal ways. On the opposite end of the spectrum, there are tremendous risks If we fail to put in place the many financial, economic, regulatory, and educational innovations that this new age demands, society may further fracture into haves and have-nots. Let's stop wasting time.

Can we really speed up progress in cognitive economics to meet the moment? To mix, match, and mangle political catch phrases: Yes, we can, but it will take a village. The changes we need require many of us to come together to challenge the status quo. With that in mind, the tenth and final chapter invites your participation in the process of acceleration. I open by sketching an ideal Accelerator for Cognitive Economics to remove current impediments to the necessary forms of team research. I close by outlining two next steps that require your involvement. The first is for you to complete a questionnaire to let me know what you think are the most important open questions and to present your own ideas for acceleration. I plan to post information on the responses I receive as well as additional material on research and institutional progress as indicated in the final chapter. The second step involves setting up a series of meet-ups that can act as stepping stones to the larger endeavor. I propose specific meet-ups of leading researchers with policymakers, business leaders, grant officers, educators and students, and concerned citizens. Interested readers are invited to volunteer for key roles in organizing these meet-ups and in other field-building initiatives. Together we can make a difference.

1.4 Why Is Cognitive Economics Challenging?

If cognitive economics is so important, why has it so long to develop, and why is there still so little urgency? One important barrier to progress is the structure of the social scientific research enterprise. Cognitive economics is a team sport. It straddles fields such as economics, psychology, law, and data science, as well as such sub-disciplines as consumer theory and producer theory. While the core models stem from economics, measurement techniques driving the research forward are largely inspired by psychology. Unfortunately, the social scientific research enterprise is slow in supporting interdisciplinary teamwork. Reasons for this are further discussed in the closing chapter of the book, where proposals are made to better support necessary team-building activities.

Equally as damaging as outdated divisions between different fields of study are the almost equally rigid divisions of each field into competing sub-disciplines. Particularly sharp are the dividing lines between theoretical economics, with its rich interplay of forces, and applied economics, which handles these interactions in the real world. These boundaries are also well-policed by loyalists on both sides of the divide. Cognitive economics requires a more integrated approach in which theory and measurement are co-created-a process of *Data Engineering for Cognitive Economics* that is covered in the paper of that name in the Journal of Economic Literature, which is complementary to the book. My upper level course book *The Science of Mistakes* (Caplin, 2023) caters to those familiar with economic modeling. But the basic ideas of cognitive economics are accessible to all. It isn't rocket science. If it were, it would be better funded.

Another barrier to cognitive economics is ideological. There are unfortunate political divisions between economics and other social sciences. Many social scientists outside economics like to view economists as politically regressive. They like to believe that the mere act of learning economics corrupts young minds into a capitalistic and selfish mindset. Many economists have the opposite problem. They like to believe that social scientists who are *not* well-trained in economics are technically deficient ideologues. They like to believe that *not* learning economics leads young minds into an anti-business and socialistic mindset. Needless to say, this form of mutual disrespect is not a good basis for building interdisciplinary bridges.

This ideological component of social science not only limits communication, it also damages our reputation. It is easy to see alignment between the openly professed political attitudes of practitioners in different disciplines and their scientific findings. While insiders can go along with this game, they cannot prevent outsiders from seeing what is going on. Ideological clubs are not ideal sources of scientific truth. Given all of this baggage, it is hardly surprising that the social sciences are having something of a crisis of replicability.

Important and upsetting as the sociological and ideological barriers to progress are, I believe that they would have been swept away long ago if not for a profound scientific challenge that stands in the way of progress. This challenge is simple to state, but hard to overcome. It was pointed out by Block and Marschak (1959) in their pioneering work. They highlighted how much easier it is to observe *what* people do than to understand *why* they do it. Mistaken decisions don't come with labels. They often relate to hopes and beliefs that do not come true. Where exactly do those show up in data?

To illustrate the key challenge, suppose I was to prefer a *sweet apple* over an *orange* over a *tart apple*. Suppose you, as the scientific observer, were to see me choose what objectively was a tart apple (albeit not so labeled) over an orange. In reality, this would be because I figured that it was likely to be sweet. But how would *you* know that? How *could* you know this was a mistake reflecting my incomplete information about how the apple tastes? Applying the standard idea that choice reveals preference, you might incorrectly infer that I like tart apples more than oranges. You might see no need even to think about mistakes I might have made or the cognitive constraints that caused them.

To provide further color on the measurement challenge, consider the analogy with classical production theory. Given its focus on production of good decisions, cognitive economics is strongly aligned with classical models of production. The difference is that in cognitive economics the inputs are various forms of *mental effort*, and the outputs are *improvements in decision quality*. Unlike the physical inputs and outputs of the classical economic model of production, mental effort and decision quality are challenging to measure. So cognitive economics faces fundamental challenges of measurement. Standard data reveals only what actually happens as a result of the choices that people make. Without a clear grasp of what decision-makers understand and what else might have happened, we struggle to assess the quality of their decisions. This leaves us in the dark about whether these choices are well-informed and how damaging any mistakes might be.

1.5 What New Forms of Data Does Cognitive Economics Need?

Given the centrality of measurement challenges, it is little wonder that the recent growth of cognitive economics has rested in large part on innovations in measurement designed to separate what people *like* from what they *know and believe*. A simple case in which this is possible suggests what has turned out to be an important path forward. Consider multiple choice tests in which one and only one of the listed answers is correct. We can safely assume that the goal of the student is to get the question right. If they get it wrong, we infer that they did not know the answer, not that they preferred to get the question wrong. If all of life was one big test and we knew all the answers, cognitive economics would be easy. It isn't and we don't. However, the case of the test provides an important lead on how to generate data for cognitive economics.

First, there is a clear analogy between how multiple choice tests are graded and how psychologists understand human perceptual limits. As detailed in Chapter 3, a typical *psychometric experiment* involves putting two different weights in peoples' hands, and asking them which is heavier. There are many features of these answers that are interesting in terms of revealing human perceptual limits. First, the answers are *stochastic* (a fancy word for random): not all answer the same way with the same weights. Second, they depend on the true *state of the world* (a fancy phrase for the truth): the left hand is picked more often when it holds the heavier weight. The third is that this *state-dependent stochastic choice* (a fancy phrase for data that is both random and depends on the state of the world) behaves lawfully: for example, the proportion of times that the correct hand is chosen turns out to depend on the ratio of weights rather than on the absolute difference in the weights. The Weber-Fechner laws of psychophysics summarize regularities across myriad realms of perception.

I illustrate repeatedly in what follows that an analogous form of measurement is of the greatest value in cognitive economics. Specifically, it sheds light on *contingent decisions*: how what people do depends on what they know and believe. I will henceforth refer to state dependent stochastic data by the acronym SDSC. In economic applications, the definition of the state of the world is far more subtle and precisely targeted definition than the object weight of an object or whether an exam answer is right or wrong. Instead, it might be the precise terms of a complex contract. It might be the actual contingent cost of insurance as a function of all possible forms of medical problems. It might be the full text of a book. It might be the historical wages of those who receive a certain college degree. It might be whether or not an image of a blood cell indicates cancer. It might be the best way of organizing workers to maximize productivity. It might be the accuracy of an AI assistant you are using in the course of your work. It might be the truth or lack of it of a seemingly factual response from ChatGPT. As will be seen, methods of generating SDSC are available in all cases and are critical to applications of cognitive economics.

The big idea is that SDSC teaches us the extent to which decision-makers understand the true state of the world. To give a simple example, when sellers of products engage in obfuscation, their hope is that they will be able to raise prices without customers noticing (Gabaix & Laibson, 2006). Evidence is to be found precisely in the lack of responsiveness of demand to changes in this price. Many other examples are provided in the book on an application by application basis.

While important, SDSC is far from the only important form of contingent data for cognitive economics. Economic surveys are increasingly in use to measure both what people believe about the future (e.g. their future earnings) and what *would happen and what they would do* in various possible future contingencies. One cannot get more directly contingent. A third insightful form of data derives from various forms of cognitive testing of value for understanding skills at work. All of these will play a significant role in the cognitive economic research that I introduce below.

We're far from finished with innovations in measurement. The drive for further such innovative measurement pervades all applications discussed in this book. Like other sciences, cognitive economics advances by developing and deploying new forms of data that are closely tied to improvements in modeling and measurement technology. We can't just rely on traditional behavioral and administrative data.

1.6 Is Cognitive Economics Different from Behavioral Economics?

The first studies of mistakes are associated with behavioral economics. These played a key role in energizing research in cognitive economics. I cover several behavioral economic topics in the book. In Chapter 2 on wealth and savings, I discuss behavioral economic research on present bias, which highlights the strong influence of immediate gratification on spending decisions, especially among the young. In Chapter 4, I explore choice procedures that lack a clear basis in rationality, yet are significant in practice. Chapters 5 through 7 delve into failures of realism, such as overconfidence, in the context of human-AI interactions. But the most direct link with behavioral economics is to be found in Chapter 3, in which I explore how changes in the presentation of information can impact understanding and choice. This analysis is inspired by the formulation of decision-making nudges in Thaler and Sunstein (2008).

In what respect is cognitive economics different from behavioral economics? The difference is that, while behavioral economics focuses on *failures* of rationality, most of the mistakes studied in cognitive economics are based on constructs that are central in models of rational learning, such as utilities and learning costs. The difference is important. Being rational does not mean being all-knowing. Students don't make mistakes on exams because they prefer low grades. Physicians don't misdiagnose patients on purpose. We don't knowingly sign contracts that waive key rights for no benefit: we simply don't realize what we've agreed to in the fine print of unreadable 50-page contracts we sign to access a website of possible interest.

Another distinction is that behavioral economics has traditionally focused less on innovations in measurement than cognitive economics. This will have to change. Given how hard it has proven to identify data to separate out preferences from beliefs, one must expect much the same when covering the wider array of forces that are allowed for in behavioral economics. There is no free lunch. Including richer forces that are hard to measure is both a blessing and a curse. The blessing is that it allows one to describe factors that may be at work in choice in a fuller manner. The curse is that it is even harder to identify measurement protocols that scientifically validate that they work in the proposed manner. I believe that the key to integrating these additional elements of richness is to identify them in data of some form. We are still in the early stages of this endeavor.

There is much to be gained as behavioral economic constructs are incorporated into cognitive economics. This is an important part of the path ahead.

References

Block, H. D., & Marschak, J. (1959). *Random orderings and stochastic theories of response* (Cowles Foundation for Research in Economics Working Paper 66). Yale University.

Caplin, A. (2023). *The science of mistakes: Lecture notes on economic data engineering*. World Scientific Press.

Gabaix, X., & Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics, 121*(2), 505–540.

CHAPTER 2

Cognitive Household Finance

2.1 The Life-Cycle Model of Savings and Wealth

Adam Smith called his book the Wealth of Nations for a reason. Wealth accumulation is key to economic development. Savings and wealth are also the focus of debates about inequity and key to policy questions such as the effects of inheritance taxes on economic growth. How people adapt their spending to short-run income fluctuations is key to the boom-bust cycle.

Questions of wealth, savings, and household finance are just as central to economic research as they are to our everyday economic lives. There are many tens of thousands of articles written about them that address a huge variety of questions. How well are we buffered against shocks, such as unexpectedly losing our jobs, having high medical expenses, etc? To what extent do we spend windfall income, e.g. from a tax refund, as opposed to saving it? What is the impact of changes in our pension plans on our overall level of savings? Why do many of us simultaneously hold low-yielding assets in banks while paying high interest on unpaid credit card debt? What role do housing assets play in our savings decisions? How important are bequest motives in determining our levels of wealth? Why do so many people avoid equities when historically they have had such high returns? Why do households with similar incomes end up with such different savings?

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_2

Economists have developed a profoundly unified viewpoint on all of the key questions in these areas of household finance, which Campbell (2006) formulated as an important and distinctive area of research. The dominant framework is the life-cycle model. We use this framework to organize our understanding of patterns of savings, wealth accumulation, portfolio choice, and more. Economists conceptualize wealth and savings using the idea that choices of how much to spend and save are motivated. The formal translation involves a utility function that summarizes our likes and dislikes that we use to trade-off what we want against what we can afford.

In the case of savings behavior, the basic trade-off relates to when to spend our resources: today or in the future. Since not all of us instantly spend everything we have (more on that later), delayed spending must provide utility. But perhaps the future does not matter as much as the present. To capture this we allow for *discounting* of utility to downgrade future relative to current consumption. We allow for many future periods with varying patterns of earnings and for different conditions for borrowing and saving. We allow for uncertainty and for risk aversion that makes it particularly undesirable to face the prospect of having very little to spend. Together these forces motivate interest in insurance. Putting this all together defines the paradigmatic life-cycle model of wealth accumulation, savings, portfolio choice, and more.

While subtle and complex in its fully formalized form, the basic ideas of the model are simple and intuitive. The life-cycle framing acknowledges that there are different motives for and patterns of borrowing, savings, and investment at different life stages. For most, the early years involve scrimping and borrowing. Those who make it onto the job ladder then face middle years in which they may have ambitions to buy homes and accumulate wealth for retirement. For most of us the later years are a time of lower income, spending down of saved assets, and worrying about health and other threats to well-being.

I am a huge fan of this classical model as the starting point for research. No reasonable critique of this approach should ignore its many virtues. I have found my personal concerns to evolve largely as this framework suggests. It is no wonder that it dominates the economic literature. Almost all research advances to this day are organized around this life-cycle framework. I cannot think of any other framework that is half as fruitful in conceptualizing wealth holding and spending motivations and their link to behavior. If you can, please let me know. Otherwise perhaps you will agree with my view that, rather than scrap the model and start again, what we need to do instead is to enrich it to account for missing cognitive factors.

2.2 The Cognitive Life-Cycle Model

As economists have studied savings and wealth, so we have uncovered intricate links with cognitive factors. I will outline three particularly interesting cases in this chapter. Of course, the importance of cognitive factors is not news to practitioners. Financial advisers play key roles in helping many of us see the future more clearly and taking appropriate protective measures. Cognitive economic research is finally bringing the worlds of theory and practice into closer contact.

What makes the life-cycle model particularly interesting from a cognitive point of view is that it buries cognitive factors in plain sight. To see why, consider the *psychology* of discounting. What exactly is the future doing in today's utility function, and what accounts for it being down-weighted relative to the present? What aspects of the future make their presence felt today? What factors determine the *presence of the future* in our utilities? Whatever they might be, they must at some level be cognitive in nature. How much and how well do we perceive our future situation? Does it depend on such cognitive factors as the extent to which we think about the future, and how we think about it? Are there methods that we can use to change our future orientation should we so wish? Can we be taught to do so?

These questions are begging to be asked, but for the most part economists have not. Exceptions that prove the rule include Gabaix and Laibson (2017), who show that noisy mental simulation can result in under-weighting outcomes farther in the future and Mani et al. (2013), who find that time preference can be affected by stress and cognitive load. But these possibilities are hard to tease out in data on savings and investment. Using methods characteristic of many sciences, economists go no deeper into phenomena than the data they have available can justify. Take a standard economic question such as the impact of changes in tax incentives for savings. Is it first order to consider how this might change attitudes to the future? And even if it is, what form of evidence would show this? To date, the answers to these questions are in the "who knows?" category. They will remain essentially unasked, hence unanswered, until cognitive economics has made a case for change. In the meantime, economists can easily justify treating them as independent of policy changes. What else can they really do?

What is true for discounting is no less true for risk aversion. In a life-cycle setting, this captures how much worse it is believed to be to have one period of very high followed by one period of very low spending relative to having a smooth path of spending with the same average. Models suggest the importance of this motivation in practice: those who are particularly averse to periods of low consumption save more for precautionary reasons. But as with discounting, economists have had little to say about cognitive aspects of risk aversion. What determines the worry that possible low future consumption induces, and how does this depend on how much one dwells on this possibility?

Overall, the status quo has involved three largely separate paths as economists, psychologists, and financial advisers think about factors at work in future orientation. Economists by and large treat the cognitive aspects of discounting as matters for psychologists rather than for themselves. While psychologists spend more time thinking about these factors, they rarely make robust connections to actual savings behavior, the gold standard method of checking that one has successfully understood discounting. Meantime many financial advisors and even financial websites offer solutions to those who would like to save more than they do. Many involve essentially cognitive behavioral interventions related to planning, etc. How do we keep our future well-being in mind and plan out a path to building up wealth for a rainy day or for high late-in-life expenses? What habits and psychological traits might help us achieve our goals in terms of work and savings? How can we learn and develop these helpful habits?

There is clear room for gains from trade. Economists are superb at working out the implications of discount rates for future behavior. They have had essentially nothing to say about how to change orientation toward the future. This is one of the gaps that cognitive economics is starting to fill in. As further motivation for enriching the life-cycle model to include behavioral factors, I discuss below three different stages of the life cycle that involve novel factors the standard model does not allow for. Before outlining these exciting areas of research, I step back to explain why they have proven so challenging and why research is still at an early stage.

2.3 Measurement Challenges

If cognitive factors play so many important roles in our attitude toward the future, why have we done so little to advance our understanding of them? As you might suspect by now, the answer comes down to challenges in measurement. For reasons that are easy to understand, the dominant form of measurement for understanding life-cycle behavior is administrative in nature. What economists do is to put together ever more granular datasets that allow us to observe rich dynamics in what people earn, how much they save, their patterns of spending, portfolio holdings, and wealth accumulation.

No one can dispute that administrative data is vitally important if we are to understand life-cycle savings and wealth accumulation. What might surprise outsiders to the field is how very hard it remains to put together comprehensive data that covers all aspects of earnings, spending, saving, and portfolio choice. You might imagine that knowing all the facts is straight forward. You would be wrong. I thought that way in my early years as an economist, in which I focused on purely theoretical matters. What I have found out since is that data limitations are profound. To this day there is not a single source with full details on earnings, spending, and asset holdings. Economists who put together the ever richer datasets that we now use deserve the greatest respect. Most of what we know derives precisely from these pioneering efforts. So hard has this proven that interest in data enrichment has been limited. If we don't even know the facts, why should we bother about enrichment?

The answer, in my view, is that administrative measures alone cannot teach us what we need to know about *why* we see the patterns that we do. Many different patterns of preferences, beliefs, and other relevant factors can explain precisely the same patterns in administrative data. It is hard to conceive of observations that could not somehow be explained by tweaking the highly flexible life-cycle framework to fit the facts. As ever, we need somehow to go beyond this factual record to understand the role of cognitive factors.

Empirical findings motivating this step into the cognitive arena are mounting. Factors that have little or no place in the standard model have been shown to change patterns of wealth accumulation. These include default savings rules and behavioral nudges. Following the work of Madrian and Shea (2001), we know now that shifts in default rules for contributing to pensions have a substantial impact on consumers' investment strategies. This is hard to square with the maintained assumption that we have thought through all possible contingencies and decided on our optimal pension contributions. If this was so, we would simply undo a default and pick out the best strategy regardless. The default would be irrelevant.

To make the case for cognitive household finance more concrete, I present three particular settings in which the interplay between financial and cognitive factors is particularly important.

1. *Out of (Self) Control*: There are suggestive indications that many, particularly younger individuals, find it hard to control the urge to spend, with damaging later consequences.
2. *Ignorant but far from Blissful*: Many in the working years do not understand financial matters well, and end up poorly prepared for retirement.
3. *Our Own Worst Enemies*? Many in the later years suffer from cognitive decline and waste resources as a result. I raise the question of how best to design housing finance markets to allow those who are concerned about cognitive decline. Addressing this head-on is one example suggesting the vast business and policy ramifications of cognitive economic research.

In all areas, we have enriched measurement to gain insight into the importance of cognitive factors. These are just a few of the myriad ways in which cognitive factors play into life-cycle decisions in the broad arena of household finance. In the arena of wealth and savings, the essential role of cognitive economics is to bridge gaps between economic, psychological, and practical approaches to future orientation.

2.4 Out of (Self) Control

An early hint that we may need to go beyond the classical life-cycle model comes from one of the great figures in the history of social science: Paul Samuelson (1937). He explored what conditions would enable a saver to make consistent decisions over time. Under what conditions would a saver be able to commit to a strategy without being tempted to change their mind later? What he showed was that this form of *time consistency* is a knife edge. Any departure from constant discounting over time, called *exponential discounting*, causes a difference of opinion between present and future self that makes plans hard to keep. Samuelson did not believe that exponential discounting was credible and assumed that we would, therefore, be forced to study commitment problems.

The next key insight came from Robert Strotz (1956), who made a specific claim about *the form* that time inconsistency would likely take. He made a case for *present bias*. He posited that the present moment often dominates our decision-making processes. The time, as they say, is always now. If now looms as large as it seems to, then people may be tempted to spend any savings sooner than planned. The challenge they might then face is how to commit themselves to not touch savings prematurely. This leads to the question of how they can effectively commit to preserving their savings. The model that describes this behavior is now known as *hyperbolic discounting*. This is a cornerstone concept in behavioral economics. David Laibson has been leading the charge in fleshing out the implications of the model for savings behavior, credit card debt, and more (Laibson, 1997).

Despite significant theoretical advances, practical challenges in measuring how much present bias actually influences behavior persist. It's not straightforward to label spending as impulsive. We simply see what is spent. For example, if someone frequently uses high-interest credit cards, it might be tempting to attribute this to poor impulse control. However, it could also be a rational response to a shortage of cash on hand, that, if not addressed, would lead to yet worse outcomes.

The larger challenge is that the traditional model with exponential discounting is highly adaptable; it can explain a wide range of behaviors, which is both advantageous and problematic. There is little that the standard model with exponential discounting*cannot* explain. This has crowded out much fruitful investigation of cognitive factors at play in the financial arena.

How is cognitive economic research advancing in this area? As always, through innovations in measurement. I have participated in research on self-control problems by modeling and measuring the gap between what people regard as their *ideal* levels of spending and their *actual* levels of spending (Ameriks et al., 2003). The headline findings are simple. For many younger people, their ideal would be to spend *less* than they believe they will. Those who are particularly prone to this hold *lower* levels of wealth, much as models of present bias suggest.

The findings deriving from our simple measures of self-control problems not only provide support for theories that allow for them, but suggest something that we had not in any way suspected. For older people, we found essentially inverse effects. Many would ideally like to spend *more* than they expect to. Those who are particularly prone to this hold *higher* levels of wealth. The implications of these findings are profound, suggesting that the nature of time inconsistency may evolve throughout one's life.

At the outset of the research, some 15 years ago, we had not foreseen that older respondents would see themselves as over-saving rather than under-saving. At the time I had no direct personal insight into how credible or important this effect might be. Time teaches us many lessons. One that I have learned in recent years is that *future bias* may indeed be more important for many older households than is *present bias*. As always in science, measurement matters.

There is much more to be done to dig into the nature of time inconsistency and how it might change over the life cycle. I see further innovations in measurement as the highest priority. Severine Toussaert has introduced a particularly important innovation in measuring awareness of *temptation* and the ability to resist it, albeit not in the area of wealth accumulation (Toussaert, 2018).

2.5 Ignorant but Far from Blissful

In the classical life-cycle model, how financially prepared we are depends on how much we discount the future, our risk aversion, and our beliefs about future income and expenditures. We are treated as far sighted and able to understand the implications of all possible decisions we might make. We then select optimally among options in light of our resource constraints.

The evidence suggests that not all is quite such plain sailing. Many arrive at retirement with essentially no assets to speak of and hence end up relying on pension income. This can result in a sudden drop off in resources available and in spending. There are ways to rationalize these findings in the classical framework: for example, some reductions in spending may be expected given the additional time that retirees can devote to cooking, the lower needs for commuting costs, business attire, etc. But there are indications that these explanations fall short.

Recent research illustrates how far off the mark is the image of an all-seeing saver making ideal decisions given well-understood constraints and rich knowledge of financial options. Lusardi and Mitchell (2014) have developed a three-question *financial literacy* instrument. Their first question measures the capacity to do a simple compound interest calculation. Their second question measures understanding of inflation, again in the context of a simple financial decision. Their third question investigates knowledge of equity markets in general and the idea of risk diversification.

One stand-out finding is that many people fail to answer even these most basic financial literacy questions correctly. A second is that those who do particularly badly on the test have accumulated little wealth. This short instrument has been shown to have a strong association with what looks like poor financial performance and financial mistakes.

This insight raises as many questions as it answers. On the one hand, if financial literacy is of such value, why are so many people illiterate? Can we be taught, or is there a deeper problem holding back learning in this area. After all, why think about finances if you see no realistic hope of accumulating wealth? And if there is such an interaction of learning with financial resources, does this set up a cycle of failure? If so, can the cycle be broken in some way? There is suggestive research by Bernheim and Garrett (2003) on the value of early life *financial education*, but a more systematic exploration is clearly warranted.

A cognitively interesting hypothesis is that *planning behaviors* play a key role both in driving literacy and in allowing people to achieve their financial goals. One can best plan for the future by thinking through what might happen later in life. The difficulties most of us have in taking this ideal long-run viewpoint open the door to cognitive factors. Some like planning more than others. Perhaps if you like planning, then you're very lucky because you're going to lay out the future better. You might think about it more. You might look more patient. You might learn more about your finances. This might be a deeper psychological characteristic driving orientation to the future.

As with all cognitive factors, planning abilities are not easy to spot in standard administrative data. They are also conceptually complex. For these reasons they are typically ignored, and what little we do know derives from innovative survey design. I have participated in research that shows that there is some truth to the idea that being a good planner helps with wealth accumulation (Ameriks et al., 2003). The research design was simple. We asked questions to pin down whether or not subjects enjoyed planning, e.g. for future vacations. We then asked questions digging into the extent to which they had planned forward in the financial realm. Not surprisingly, those who liked planning for vacations had also planned more for their financial futures. Tellingly, there was a pass-through effect on wealth accumulation. Those with a high *propensity to plan* not only were further ahead in their financial planning, but also appeared to accumulate higher wealth as a result.

The results, while suggestive, leave many unanswered questions. There is far more to be understood about planning propensities. We do not know the channel through which planning relates to wealth accumulation. If it relates to bringing the future to mind, it suggests the need for new models that interact with attention and preferences. Another obvious follow-up question is whether one can teach people to be better at, or better yet to enjoy, financial planning as a process (I know the answer to this question at the personal level and find it discouraging). How do we find out whether you can actually get somebody to plan more and enjoy it? And whether they then become good at accumulating wealth and pursuing long-run career goals? Other related questions jump out. What other goals does being a skilled planner promote? After all, it would seem to be essential not only for financial decisions but also for other aspects of life, including preparing for physical and cognitive decline of which all of a certain age live in fear, and to which I now turn.

2.6 Our Own Worst Enemies?

There are many fascinating and under-studied aspects of late-in-life financial behavior. First, and most strikingly, the basic pattern predicted in the life-cycle model does not play out straightforwardly in practice. The idea that many save during the working years to spend down in retirement is evident only in the very final years. Before that, those with moderate wealth generally cling onto it (Poterba et al., 2011).

Unfortunately, there is one really significant late-in-life risk that remains essentially impossible to prepare for cognitive decline. Cognitive decline makes older Americans less capable in terms of financial decision-making and vulnerable to financial fraud. A high proportion of Americans 85 years or older have dementia. Vulnerability to manipulation and even out-and-out fraud is much higher for those who are cognitively declined. This has pervasive financial implications across the income and wealth spectrum. Social security is highly vulnerable to fraud. It is relatively easy to send out checks, but far harder to know what they are being spent on and to whose benefit. A major part of this worry may relate to financial exploitation of elders by strangers, be it through internet scams, false friendships, etc. But there are also many nightmare stories of relatives essentially stealing from their elders.

For those with significant resources, planning for possible later decline is key. But there has been essentially no research on what asset holders in the pre-decline phase are doing to prepare for possible later cognitive decline. Passing control to a loved one at some relatively early point in the process of decline may be the strategy many of us have in the back of our minds. Even here there are many unknowns. One obvious question is how trustworthy an agent such as a partner or child might be. There may also be reasonable differences in priorities. What responsible parent of young children would like to see the family fortune dissipate in supremely high bills for taking care of a cognitively declined elder?

By now, you may not be surprised to find out that economists know next to nothing about topics related to cognitive decline and its impacts on spending and wealth. They are horribly under-researched due at least in part to challenges in measurement (as you will see I suspect that there are other less savory reasons). Cognitive decline is not even measured in standard administrative data. This is even more true for worries about future decline, and as for worries about recognizing future cognitive decline, that is three steps beyond administrative data and consequently has by and large been ignored.

Understanding this massive gap in knowledge, I have participated in survey-based research with a panel of older U.S. wealth-holders to get high-level qualitative and quantitative insights (Ameriks et al., 2023). We explored worries panel members might have about cognitive decline and steps they might take to assuage those worries.

First, the bad news. The vast majority were aware of the chance that they might in future decline, and understood that they might survive a significant time, five years or more, in a state of decline. Now the good news. Most in our particular survey sample seemed relatively sanguine about the quality of the third party they would hand over to in case of decline. Around 70% reported that a child would be the most likely agent and 10% a sibling, with the remaining essentially equally divided between a trustee or institution and a broad "other"category. The great majority of respondents reported that they believed the agent would be excellent or very good at taking care of them were they to hand over control of assets.

The next obvious question is whether or not respondents believe that they would recognize their own decline in a timely manner. Timing this transfer well is complex. Most indicated a desire to stay in charge in the early stages of decline, but not in the later stages. Picking the ideal time involves a complex calculus balancing rational discomfort at the idea of handing over financial control against the possibility of hanging on too long, thereby damaging not only ourselves but also our loved ones as we squander resources. Quite a conundrum.

The challenges go even deeper than this. It is one thing to decide ahead of time the precise mental state one would need to be in to hand over control. It is quite another to be able to recognize that one has arrived in that state and to follow through on what had in the past seemed like the ideal timing. One respondent during an online chat summarized this possible source of worry succinctly in the context of handing in a driver's license:

> My mom, who is very old, was refused renewal of her driver's license because she failed the vision test. Her response was to sue the DMV for incompetence. I sincerely hope to have self-driving cars before I get to that stage.

Is this a rare story, or are concerns with unrecognized decline widespread? Evidence suggests the latter. Mazzona and Peracchi (2020) estimate that cognitive decline, of which subjects are unaware, results in 10% loss of wealth among wealthy stockholders.

The research that I have participated in provides more color on the extent to which this possibility is recognized ahead of time. Many of us appear to be aware that our future selves may turn out to be our worst enemies. Specifically, there are worries that we may delay the handover of control for too long. While most of us know we might decline and that we would ideally hand over control to a loved one at that point, we also know that we might not recognize that we have declined and fail to hand over in time. This was precisely what worried most of our respondents. While most had access to trusted agents who would have their best interests at heart, they did not trust *themselves* to recognize their own cognitive decline and hand over the financial reins in timely fashion.

The interaction between cognitive decline and aging is not only of academic interest but also of profound importance for public policy. We need to have better ways of tracking *financial abuse* of elders related to social security and private pensions. This is technologically simple and it is shocking that nothing is being done. We need to develop methods to flag obviously suspicious transactions as prompting an interaction to check that there is no fraud. In addition to wiring massive amounts of funds to previously unknown parties, there may be other cases that would warrant at least a check-in. A related issue is what to do if there are indeed good reasons to worry: Might it be time to inform a loved one, and might this too be agreed ahead of time?

What about the possibility of introducing certain agreed checks on cognition? Possibilities could be greatly enhanced with prior agreement about some forms of cognitive testing. Of course, better information alone may or may not be enough to convince us to hand over control. How can we know if, facing cognitive decline that we only partially recognize, we would go along with an earlier agreed strategy of handing over partial control to a once-trusted agent?

One sign that a subject has been under-researched is when simple questions hit a strong nerve. That is the case with the research on financial aspects of cognitive decline in Ameriks et al. (2023). It makes it clear that issues of cognitive decline are pervasive and need to be understood far better than they have been. In addition to the measurement challenges that I have repeatedly highlighted, a second reason they have been studied so little is a result of the field divisions noted in the introduction (this is not the promised unsavory reason: that is still to come). I have outlined above how economists, who traditionally focus only on administrative data, have essentially ignored the impact of cognitive decline. While there are important and honorable exceptions, gerontologists, psychologists, and others interested in how aging impacts cognitive faculties know next to nothing about economic research. The division between disciplines is extremely constraining, as are so many of the boundaries that have held back progress in cognitive economics.

Now for yet more armchair psychology. What do I have in mind as the unsavory reason for our failure to address the challenges of cognitive decline? To give you a hint, it is related to what humans do and do not like thinking about. Ask *yourself* how much you enjoy thinking about cognitive decline, either for you or a loved one. While you may or may not be psychologically drawn to it, the odds that those are fun thoughts seem low. The young are attractive and dominate popular culture. The old less so. The cognitively declined old even less so. Thinking about one's own possible path of cognitive decline and how to guard against it is not fun.

Failure to think enough about cognitive decline is not just an academic thing. Think of the flow of venture capital. Young venture capitalists are far more interested in seeking perpetual youth than in addressing the problems of those born before the promised *singularity* that will bring them eternal youthful life. Perhaps this explains to some degree why so little effort has been devoted to innovations related to cognitive decline. So it is not only researchers and policymakers who need to up their game. The private sector should focus on innovations that are cognitively robust.

There are some obvious market openings. A particularly important role might be played by innovations related to owner-occupied housing. Long-term care in a facility is extremely expensive and getting ever more so. It is also aversive for many. Many will want to receive support services that might become necessary without immediately moving into a care facility. There are few who want to be burdens on their children, and fewer still for whom this is agreeable to the children themselves. It seems highly credible, if under-researched, that most elders would prefer to age in place unless their physical or cognitive decline makes this impossible. However, there are great financial challenges, particularly as physical and cognitive decline involve medical and care expenses even for those who stay in their own homes. This raises the question of how to finance staying in the home.

Cognitive economics suggests some possible paths forward for those who own their homes. In essence, we need to introduce market mechanisms to provide funding out of the owner-occupied home itself, since this is for many the most important asset. Here some early ideas on markets in housing equity seem relevant (Caplin et al., 1997). The idea is to create markets for shared ownership of housing equity that allow partial transfer of ownership. I know of at least one ongoing effort to develop just such markets and see them as having huge potential to help those who wish to stay in the home without becoming burdens on loved ones. For any entrepreneurs who plan on aging the old-fashioned way, markets are ready and waiting to be developed.

References

Ameriks, J., Caplin, A., & Leahy, J. (2003). Wealth accumulation and the propensity to plan. *Quarterly Journal of Economics, 118*(3), 1007–1047.

Ameriks, J., Caplin, A., Lee, M., Shapiro, M. D., & Tonetti, C. (2023). Cognitive decline, limited awareness, imperfect agency, and financial well-being. *American Economic Review: Insights, 5*(1), 125–140.

Bernheim, B. D., & Garrett, D. M. (2003). The effects of financial education in the workplace: Evidence from a survey of households. *Journal of public Economics, 87*(7–8), 1487–1519.

Campbell, J. (2006). Household finance. *Journal of Finance, 61*(4), 1553–1604.

Caplin, A., Chan, S., Freeman, C., & Tracy, J. (1997). *Housing partnerships*. Cambridge, MA: MIT Press.

Gabaix, X., & Laibson, D. (2017). *Myopia and discounting* (National Bureau of Economic Research Working Paper 23254).

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics, 112*(2), 443–478.

Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of economic literature, 52*(1), 5–44.

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science, 341*(6149), 976–980.

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *Quarterly Journal of Economics, 116*(4), 1149–1187.

Mazzonna, F., & Peracchi, F. (2020). Are older people aware of their cognitive decline? Misperception and financial decision making. *Journal of Political Economy, 132*(6), 1793–1830.

Poterba, J., Venti, S., & Wise, D. (2011). The composition and drawdown of wealth in retirement. *Journal of Economic Perspectives, 25*(4), 95–118.

Samuelson, P. A. (1937). A note on measurement of utility. *The review of economic studies, 4*(2), 155–161.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies, 23*, 165–80.

Toussaert, S. (2018). Eliciting temptation and self-control through menu choices: A lab experiment. *Econometrica, 86*(3), 859–889.

CHAPTER 3

Measuring and Minimizing Mistakes

Let me reiterate the core measurement challenge of cognitive economics. Basic economic reasoning implies that the choice of one option over another reveals a preference for it. Yet when information is imperfect and mistakes are made, this simple equivalence breaks down. Standard data *do not and cannot* identify mistakes. What we see in standard choice data is what people choose. This reflects both what they like and what they understand. We cannot infer much about the quality of the choices that they make without separating out these factors. To observe is easier than to explain. Choice alone provides little evidence of mistakes that are being made and any damage they may be doing.

Let's translate this challenge to the field. A scientific observer of school choice cannot *prove* that families applying to 1000 schools in New York are profoundly confused. Given this, it is perhaps no wonder that most economists studying matching mechanisms assume that information is complete. Not a happy state of affairs but understandable, given that mistakes do not come with labels.

If you think lack of realism is just an economics thing, think again. Most of the examples in this chapter are from the legal and regulatory realms. Think *informed consent*. In many countries, the U.S. and the UK included, basic contract law enshrines the obligation on those who sign contracts to understand them, without any checks on how comprehensible they are. What did you really sign up to last time you clicked *Agree* before entering a website? Was it a well-considered decision reflective of

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_3

your priorities? Or was it instead a result of the incomprehensible legalese you would have had to read to take the decision seriously? Knowing the latter to be common, regulators understandably step in to try to enforce comprehensibility using mandates. But they have no system in place to check that these mandates improve effective understanding or reducing mistakes.

Before getting into matters legal and regulatory, I open the section by outlining field studies in which mistakes are identifiable and can be reduced by changing presentation. This method of changing presentation to change behavior links with the literature on "nudges" (Thaler & Sunstein, 2009) and design of *choice architectures*. The additional challenge in the cognitive economic approach is to confirm not that *social goals* are met, such as raising savings rates, but that *private goals* are better met. This is notoriously difficult in complex choices in which all options have pros and cons.

My discussion of legal and regulatory matters focuses on how little we understand about effective communication. I outline a field study in Mexican labor courts which reveals how socially important it is for us to take this challenge more seriously. I report on an intervention in which clearer presentation of information demonstrably *raises the quality of justice*. So clear-cut are the findings that they have been incorporated into ongoing changes in protocols in cases of legal conflict in Mexican labor courts. The U.S. and other countries would do well to follow suit.

I close the chapter by outlining new cognitive economic methods for measuring how effectively complex information is communicated, with an important application to *online privacy disclosures*. Refining such methods is a necessary step if we are to revamp standard methods of communication and the regulations surrounding them to be effective in the era of information overload. With further research, corresponding methods could be put in place worldwide. As so often, the innovations relate to effective measurement. We need methods to replace subjective plausibility checks in the legal and regulatory communities. What we need are new cognitive economic measures of effective communication for complex environments.

An open question is the extent to which the private sector will evolve to a transparency norm in which unclear presentation is punished in the market. I suspect there will be profound limits to this self-policing activity, making effective regulation critical. I close the section by outlining ongoing research on the design of such regulations, with a particular focus on *privacy disclosures*.

3.1 Mistakes in Complex Choices

We often have to choose among complex options, and it is unreasonable to imagine that these choices are fully informed. Yet it is hard to prove otherwise when most options have pros and cons, as is typically the case. To date, the best evidence has been found in work settings rather than in once-off consumer choices. Not only are tasks repeated often, but good ex post measures of performance are often available. Even here one often needs to be imaginative in putting data together.

One valuable method involves tracking *the process of search*. While one can debate what is learned when information *is* accessed, failure to access reveals profound limits in understanding. This has proven useful in resume studies. Bertrand and Mullainathan (2004) conduct resume studies by changing the signaled ethnicity and gender of applicants in fictitious resumes. They find a strong disparity in callbacks for a given quality of resume depending on the signaled ethnicity and gender. One channel that they suggest is cognitive, for example, reading no further in a c.v. when an employer identifies a name as signaling a negatively stereotyped group. This is studied formally by Bartoš et al. (2016) who model an interesting distinction between markets involving lemon-dropping (e.g. rejecting one of a small number of applicants for rental housing) as opposed to cherry-picking (accepting one of many applicants for a job). By cleverly designing the search interface, they show that negatively stereotyped minority names indeed reduce employers' effort to inspect resumes, while increasing this effort in rental applications. This confirms the important difference that the theory suggests between lemon-dropping and cherry-picking aspects of prejudiced decisions.

Another important method of revealing limits on understanding is to explicitly provide relevant information. In a field study of experienced Indonesian seaweed farmers, Hanna et al. (2014) identify limits on understanding of the underlying production technology. Farmers submerge strands of seaweed (or "pods") into the ocean and productivity is affected by many factors, not all of which are evident to farmers even after many years of practice. As so often, there are scientific studies that reveal aspects of productivity that, while important, are not top of mind for the farmers. The study identified one such factor that impacted productivity yet appeared to have escaped the farmers' notice. Most do not have a clear opinion on the importance of pod size despite its objective importance. The failure to optimize pod size meaningfully reduces farmers' output and income. In confirmation of this specific channel, an information intervention focused on pod size leads to improvements in productivity.

There is one work setting in which mistakes can be identified cleanly: sporting calls. Here ex post review can produce a perfect 50-50 vision. Archsmith et al. (2021) exploit high-frequency data on the accuracy of umpires' calls in Major League Baseball. They develop a model of the importance of each call based on the chance that a mistake would significantly change the balance of play. They find that umpires make fewer mistakes in higher stakes decisions and that the expectation of higher stakes in upcoming decisions leads to more errors in current decisions, as if the umpires are deliberately saving cognitive resources.

A recent study takes this in an important new direction involving human-AI interactions. Almog et al. (2024) study tennis, in which a human umpire makes the call, which can get overruled by AI (Hawkeye) when wrong. They estimate that umpires lower their overall mistake rate based on the subjective costs of being overruled by AI. They also shift the nature of the mistakes they make from calling a ball out when in to calling a ball in when out Almog et al. use cognitive economic methods to quantify how much umpires change priorities under AI oversight.

3.2 Appeals Courts and Error Correction

It is widely understood that a major role of the legal system is to reduce mistakes and that cognitive constraints make it difficult to achieve the ideal of error-free decision-making. Stephenson (2010), in particular, highlights features of institutional design that may play a key role in error reduction. Yet this has been little studied in the field. Legal scholars have borrowed from economists the idea that utility matters to judicial outcomes. Unfortunately, they have if anything done us one better in their neglect of cognitive constraints.

What forms of evidence can be used by those wishing to understand legal mistakes? An important path forward relates to successful appeals. Unlike in sporting calls, hindsight in legal situations is not 50–50. However, some errors are revealed when verdicts are successfully appealed. In fact, allowing such appeals is a key incentive the system provides for limiting mistakes in the first place. Typically, appeals impose time and reputational costs on the responsible parties. As highlighted by Shavell (1995) and Cameron and Kornhauser (2006), a key role of imposing such costs of being successfully appealed is to ensure that reasonable effort is put into getting the verdict correct first time.

Given that the central role of appeals is to reduce mistakes, one might expect them to be well-studied. What factors determine the proportion of decisions that are successfully appealed? Are there particular mistakes that are disproportionately common, and why might that be? How can one reduce the volume of such appeals and get rid of particularly problematic patterns? Unfortunately, there are few if any studies that address these pressing questions. It is disappointing that, for the most part, these issues of error reduction have been ignored by legal scholars. Perhaps it is easier to focus on ideology than it is to innovate in measurement as would be required to separate ignorance from intent. That says more about the structure of the legal research enterprise than about the achievement of justice.

The deep challenge, as always in cognitive economics, is one of measurement. Environments suitable for empirical field study of judicial mistakes would ideally need to satisfy a number of conditions. The researcher needs access to detailed data on the evidence presented to the judge in a large number of cases, the evidence *considered* by that judge, and the final judicial disposition of the case. Ideally, the context would be a court that has to deal with many similar and legally simple cases, the correct disposition of which can be determined based on a relatively limited number of observable variables. Finally, the researcher should be able to observe sufficiently frequent overturns on appeal, the main error correction mechanism available in the legal system. I now turn to a case study with just these features.

3.3 Indexing of Complex Information Reduces Injustice

Caplin et al. (2024) present the first field study that directly confirms the impact of cognitive constraints on the achievement of justice and of simple cognitively informed interventions to raise its quality. The context is a Mexican labor court, which provides a uniquely suitable environment for a study of this type. The vast majority of the cases dealt with by the court are *unfair dismissal* lawsuits, and we use only these cases for the study. If the employee proves that an employment relationship existed, the burden is then on the employer to prove that dismissal was fair. In trying to establish this, most employers either deny that the plaintiff was ever an employee; claim that the plaintiff left the job or resigned voluntarily; or make an offer to reinstate the worker. Submissions are also fairly standard. In a typical case, a worker might allege being illegally dismissed after a given number of years of service, and provide a written contract or salary receipts. The firm might present its payroll accounts for the relevant period to demonstrate that it never paid anyone with that name. Alternatively, it might acknowledge the employment relationship and produce a resignation letter signed by the worker. Authenticity of all these documents can be objected to by the parties, with corresponding expert opinions being attached to the case file as additional evidence.

Once cases are disposed by the labor court they can be appealed to a higher court, alleging violation of a party's constitutional due process rights. Of particular value in interpreting how mistakes are changed by clear presentation in the Mexican case is that judges are charged not only with *reaching* a verdict but also with *justifying* it based on identified *legal facts* extracted from the case file on which it is based. What allows a verdict to be appealed are identifiable mistakes in this justification.

In our field study, we were able to observe all appeals that were granted, since such cases would return to the same court for a revised disposition. An advantage of our case study, and indeed part of the reason that it was conducted, is the high proportion of granted appeals. In most courts, successful appeals are relatively infrequent, and studying the quality of first instance decisions using overturns on appeals is empirically challenging because of the sparseness of data. Yet in our case, over 30% of the original decisions in this court were appealed, and in about half of these appeals were successful in full or in part. This may indicate relatively easy access to appeals, but also low quality of the first instance decisions.

As part of a major investigation into the efficiency of Mexican labor courts, we were invited to test out methods of improving the quality of judicial decisions. We implemented a randomized intervention directly varying the presentation of the cases to the judges. Crucially, at the time of our study the court did not conduct oral trials; written evidence admitted into the case file was the sole basis for the decision produced by court officials, who were not themselves present in any oral hearings or depositions. We obtained access to complete case files, putting us in the same position as the judges themselves.

The case files themselves are lengthy and intricate. They might be on the scale of a small book, which has to be reviewed under great time pressure (the average time taken to reach judgment is approximately two days). To their great credit, the judges and the court were interested in reducing the high rate of granted appeals. They allowed the research team to test a cognitively inspired remedy. We were seriously constrained in that the files themselves could not be changed in any way, since they are faithful records of the entire legal proceedings. So what to do to help with comprehension?

We used a particularly simple and replicable strategy. While we could not alter the file in any way, the court allowed us to *add an index* to half of the files at random. For each case file we produced a 2-page summary, which included legal claims, facts alleged, and the first 5 items of evidence of each of three major types (documents, testimonies, and depositions) admitted into the case file, with page numbers to indicate where each evidence item could be found. The summaries were only provided to the judges for a randomly selected subset of cases, with the others forming a control group. There were no statistically significant differences across treatment and control groups in the number of items of evidence submitted by plaintiffs and defendants, the wage claimed by the plaintiff, or the percentage of case files in the top quartile of items of evidence of plaintiffs and defendants.

The judges were aware that both the treatment and the control group files were part of the experiment and that the summaries existed for all case files. The experiment was explicitly described to them as part of a study designed to evaluate possible procedural improvements in the court. While the experiment required close coordination with the draft decisions office of the court, the indexes were worked on by research assistants, most of whom were either upper level undergraduate law students or very recent graduates. The overall cost of our intervention was only a very small part of the cost of even the initial disposition of the case by the judge.

Our access continued even after the initial judgements. Following these judgments we were able to trace the files to determine whether or not the original decisions were later overturned on appeal. The consequent coding of the written opinions was likewise done for all the cases in both the treatment and the control groups.

The results of our field intervention are now in and show that it produced a significant reduction in the number of granted appeals. As you might expect, adding an index appears to allow judges to better identify key items of evidence and to rationalize their decisions without making clear legal errors that define the grounds for a successful appeal. Our intervention reduced overturn on appeal from 16% to 10.6%. Given that our intervention did not affect the content of the case file, but only the ease of accessing this information, we interpret this result as demonstrating that simplifying the attentional task the judge is facing by itself improves the quality of judicial decisions.

There is a big difference in the treatment effect of our cognitive intervention between objectively *complex* and *simple* cases. What defines a case as complex is that the existence of a labor relationship is acknowledged by both sides. What the judge must decide in such cases is the nature of its termination, on which both parties present evidence. What defines a case as simple is that the defendant denies the existence of a labor relationship outright and, therefore, presents no evidence.

It turns out that our treatment effect is *entirely concentrated on the more complex cases*, where the treatment halves the rate at which cases are overturned on appeal. In stark contrast, there is no treatment effect for the simpler cases. This interaction between complexity and treatment effect of simplification is striking and likely to be of great general applicability. In that sense, our treatment supports the recent focus in cognitive economics on modeling complexity (Oprea, 2020; Puri, 2018).

Being able to compare the original case files and the written opinions of the judge allows us to further explore the impact of our treatment on the content of the opinion. Thus, we can find out whether the evidence presented at trial is actually mentioned in the judicial opinion. There is a clear result: our treatment made omissions more common in the more legally complex cases in which we find the treatment effect. Our results may be indicative of the treatment effect working through shortening judges' decisions, and in particular, through them omitting the discussion of evidence that may not be relevant to their theory of the case.

The beneficial impact of our intervention not only for the quality of justice in the cases considered, but for the functioning of the legal system, is clear. In fact it was clear to the court itself. Due to the research team's strong ties to the labor authorities, we were invited to present the results of our field interventions, including this experiment. We were then able to propose parts of the new labor lawsuit procedures. The new federal labor law, passed in 2019, makes major changes to the lawsuit process. Plaintiffs now must provide, as part of their initial filing, a list of items of evidence they will present at trial, with a justification for each. In the defendants' answer they must also provide any rebuttal evidence and its justification. Trials are now oral and presided over directly by the same judge who is to write the decision. Finally, courts are obliged to have electronic case management systems, including an electronic mailbox in which parties can see all the dispositions in their case in real time and can receive formal notifications from the court. We are now designing follow-up fieldwork on the *arbitration* process and its economic ramifications. In this we are working with Daniel Chen, who has participated in pioneering research redesigning judicial practices in Kenya to clear economic benefit (Chemin et al., 2023).

A simple illustrative search model illustrates the central cognitive mechanism at work. Sight unseen, it assumes that guilt is the more likely verdict. This reflects the general pro-worker structure of the Mexican labor law. The guilt rests on two forms of evidence. It requires both that there was a valid employment relationship, and that it was not terminated appropriately. To match this we model the facts of the case as defined by two independent items of evidence related, respectively, to the existence of an employment relationship and the process of termination. Each can take be of type g and i, with g the type that is necessary for guilt to be established in either case. Hence for the employer to be guilty requires items of evidence to be g: the firm is innocent if either there was no employment relationship or there was but it was fairly terminated. The goal of the judge is to minimize the expected number of mistakes with utility losses normalized to 1 of making a mistake.

To be able to write an opinion, the judge must view at least one of the two items of evidence. The only question for the judge is whether or not to observe the second item of evidence after looking at the first. To capture idiosyncratic features of each case file, suppose that there is a random cost of undertaking the second search that is known when deciding whether or not to observe the second piece of evidence. If realized costs are at the top end of this range, search is certainly not worthwhile. If it is at the base of this range, it certainly is. Hence whether or not the second item is searched is random.

The way to think of our treatment of adding an index is that it makes it possible for the judge to identify up front the more important information, and to use this to determine search order. In our simple formulation, this is the one that is more likely to reveal innocence. It is easy to show that it is optimal to search for this more important item first once the index allows it to be identified. Relative to unguided search without the index, this both reduces how often mistakes are made and shortens the expected number of items listed, just as we found.

The fact that there is such a simple and transparent rationalization for our findings suggests strongly that we did not find the only case in which cognitive factors interfere with the administration of justice. What is alarming is only how *little* research has been designed to improve the quality of justice by improving communication at all stages. All we did was to add an index up front to guide the reader in how to assign their cognitive effort. In essence, our model captures the obvious idea that search is more efficient with an index. This makes it crystal clear that the main channel of impact on decisions is attentional. It is past time to put analogous methods to work more broadly.

3.4 The Duty to Understand

There are countless settings in which cognitive constraints give rise to decision-making mistakes that might be reduced by simpler presentation. Consumer contracts are obvious cases in point. In many countries, the law makes it the duty of those who sign contracts to understand them. The application of this duty is especially controversial in the context of consumer contracts, which consumers generally do not read and are, in fact, very hard to understand.

In an article aptly entitled "The Duty to Read the Unreadable", Benoliel and Becher (2019) patiently explain the duty to understand doctrine, which holds contracting parties responsible for the written terms of their contracts, whether or not they actually read them. In making their case for changing this doctrine, they undertake readability research on "sign-in-wrap contracts". These are agreements that an online website requires users to accept before they can use its services. These agreements combine the process of signing up for a website with agreeing to its terms and conditions. Typical clauses that are included are: an intellectual property clause; a prohibited use clause; a modification clause; a termination clause; a limitation of liability clause; a disclaimer clause; a class action waiver clause; an arbitration clause; a forum selection clause, which establishes; a governing law clause; and a time bar clause. Who knew?

Sign-in-wrap agreements and the duty to read them have already been at the forefront of various legal battles and courts routinely apply the doctrine. To see how realistic is this demand, Benoliel and Becher apply well-established linguistic readability tests to the five hundred most popular websites in the U.S. that use sign-in-wrap agreements, which include Google, Facebook, and Amazon. Readability is assessed using a well-established grade-level test. The maximum recommended score for consumer-oriented texts is 8.0. If the goal was to have consumers understand their options, this would be the appropriate standard. Do you think it is being met?

As you likely guessed, the findings are not pretty. The test is failed in a dramatic way. The median grade level of sign-in-wrap contracts is 14.9. This is comparable to the usual score of articles in academic journals. When the Duty to Read meets the Intent to be Unreadable, you know which wins.

Noting that many contracts that we routinely sign are unreadable and unread, Benoliel and Becher propose a regulatory remedy: imposing a general readability *duty* on consumer contract drafters. Under the suggested readability duty, drafters would be required to provide contracts that consumers can easily understand applying the grade-level readability test.

At first hearing, this sounds good. But there is a gap in their argument. The value of their proposal rests on mandates being effective. After all these decades, one might have expected regulators to provide proof that their methods are effective. Unfortunately, that is far from being the case. No regulatory tool has had such an expansive application in clarifying information as the *mandated disclosure* statement. Mandated disclosures appear everywhere from credit card statements to cafes. Yet evidence of their effectiveness is very thin, as stressed by Ben-Shahar and Schneider (2017).

As always, the challenges of measurement are holding back the application. What helped our work in the judicial setting was the central role that appeals courts play in the identification and rectification of errors. There is no equivalent in everyday contract choice. Mistakes are doubtless common, but which decisions are and which are not mistaken is hard, if not impossible, to judge using data on choices alone. The larger issue is that without any method of measurement, it is traditional worldwide to ignore the damage done by needlessly complex communication and needlessly onerous rules. Yet again, the fantasy of frictionless understanding rears its ugly head. If we do not have methods of measuring damage done, we traditionally ignore this damage. We really do need to up our measurement game in this regard.

3.5 Bloom's Taxonomy and the Goal of Communication

The challenge for regulators is to somehow test out methods of simplification in settings in which errors *can* be measured. The open research questions, as always, are the *how to* questions. How can we design *measurement protocols* to reveal how well information critical to decisions is communicated? How can we design regulatory protocols that can be verifiably shown to improve such communication? Current regulatory procedures involve a priori decisions on the form of communication that is to be deemed simple. These are often prima facie credible, for example, insisting on uniform presentation of annual percentage interest rates (APRs) for credit products, designing uniform nutritional labels for food items, and writing important information in large capital letters. What has been missing to date are procedures for validating these methods as producing positive outcomes for consumers. In what follows I outline methods that are currently being developed with precisely this goal in mind. Taking this next step is important for building trust in the regulatory process itself. Regulators increasingly recognize the need to set higher standards for themselves, as witness White House Executive Order 14058 (2021) on rebuilding trust in government.

The first question that needs to be answered concerns the goal of regulations in areas such as online contractual disclosure. Here it is of value to draw on Bloom's famous taxonomy of educational goals (Bloom, 1956). The first level of expertise in this taxonomy is *raw knowledge of facts*. The second level is *comprehension*, which involves some form of intelligent interpretation. The third is *application*, in which understanding is shown by the accurate incorporation of knowledge in concrete settings of importance.

In terms of communication, it seems that it is essential to have at least the third of Bloom's levels of understanding as a goal. The key goal of communication, and of disclosure mandates, is presumably to help consumers make better informed decisions, rather that to endow them with a list of facts and statistics whose translation into decision relevance is beyond them. For example, in the case of contractual obligations, the question is whether or not the clauses are presented in such a manner that decision-makers can accurately reflect their priorities in deciding whether to accept, decline, or effectively select among various options. The goal must be to *allow those with specific priorities to best identify what is right for them.*

It is important to note just how far short of this ideal current methods of presentation and of regulatory measurement fall. As noted, the method that is most common in practice is subjective plausibility in the minds of the designers and regulators. Those looking to validate regulatory measures sometimes go beyond this by using standardized grade-level comprehension checks. But even this falls far short of ideal. Designing a test to see if these facts are internalized is only a level 1 check. Getting consumers to identify facts is the first level in this taxonomy, while the second level is for them to understand these facts properly. This cannot be taken for granted. For example consumers who can remember the APR on a loan but have little idea what it means in terms of dollars and cents achieve level 1, but fall at level 2. But what we really want to know is one level higher. The third level in Bloom's taxonomy is being able to apply understanding effectively in decision-making. Grade-level tests and factual quizzes address at best the first two levels, but miss out on this crucial third level. In what follows I introduce protocols that are currently being designed to test achievement of this key goal.

In order to observe whether level 3 understanding is achieved, one has to be able to observe whether an individual makes an optimal choice or a mistake. As always, the challenge is that standard choice data (e.g. opting in or out of a service) reflects both what consumers like and what they understand. Naturally, this makes their level of understanding extremely difficult to measure via choice alone. Whether or not a choice was made because it was desired is very challenging to assess.

3.6 Cognitive Economics of Effective Communication

As I have noted often in this book, it is very hard in the open field to identify mistakes induced by a given method of communication. For that reason what researchers need is an experimental wind tunnel. We need a method to compare presentation protocols for consumers whose priorities we know so that we can precisely measure patterns in the resulting mistakes and how they vary across different presentation protocols. Cognitive economic methods provide a precise method of doing just this.

Before introducing the new methods cognitive economists are developing, it is helpful to go in a bit more depth into the psychometric tradition. This dates back nearly two hundred years to the work of Ernst Weber (1834) and the introduction of *psychometric curves*. These are the most basic constructs in many areas of psychology. They measure the limits of human perceptual ability. For example, they can be used to characterize the proportion of times that the heavier of two hand-held weights is correctly identified as a function of the difference in weight.

Figure 3.1a draws a typical psychometric curve. The horizontal axis in the figure measures the objective difference in weight, d, between left and right hand, which the experimenter knows but the experimental subject can only approximate. The vertical axis measures the probability that the object in the left rather than the right hand is chosen by subjects as heavier.

In a simple task such as this, it is typical for there to be no clear bias: if the weights are in fact equal in both hands, then both hands are equally likely to be *perceived* as holding the heavier weight. That is reflected in the figure by the curve passing through the y-axis at probability 0.5 when $d = 0$. For values of $d > 0$, the left hand holds the heavier weight and is correctly perceived as holding it more often than not. Reflecting this, the point that is marked in the figure shows that when the difference is $d > 0$, the heavier weight is identified with probability $P(d) > 0.5$. As the difference in weight gets larger, so the heavier weight is correctly identified more and more often, ultimately heading to a point where the difference is large enough that subjects always make the correct choice. The figure also illustrates symmetry in errors between hands. For any given difference in weight, it is equally likely to be identified correctly whether it is in the left or the right hand.

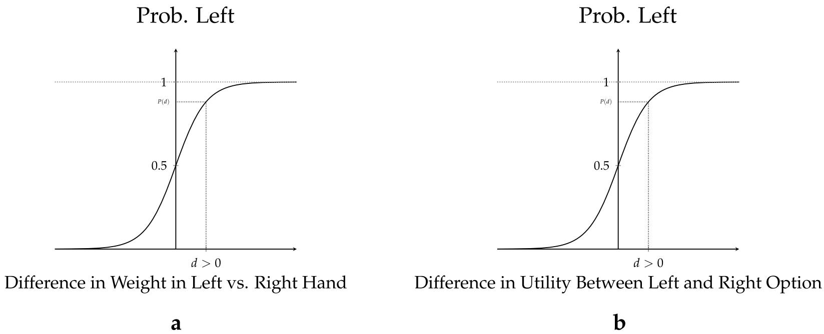

Fig. 3.1 **a** Psychometric curve. **b** Cognitive economic curve

How does this relate to the challenge of measuring how well decision-makers understand the options they face? The link is indirect but precise. Consider a variation on the perceptual theme in which what has to be identified is not which of two weights is heavier, but rather which of two complex options, again placed on the left or the right, better meets consumer needs. To be precise, we consider an experimental design in which the researcher has controlled the set up so that they in fact know not only which is better for each consumer, but also by how much in terms of utility.

In such an ideal case, what we might expect conceptually is very much in line with what we see in the case of weight. On the whole, consumers will be more likely to pick the option that is better for them, and they will be ever more likely to as the stakes in the decision get larger. So we would expect most to make the correct choice given their values, but also for some to make mistakes. If that is the case, then a curve such as that drawn in Fig. 3.1(a) might serve to record the proportion of mistakes as a function of the true difference in value.

Figure 3.1(b) draws just such a figure. The main difference is in the name: we call this dataset a cognitive economic curve rather than a psychometric curve. Correspondingly we relabel the horizontal axis as measuring the difference in utility between the options on the left and on the right, with the vertical axis measuring the probability that the option on the left is chosen. Conceptually the cognitive economic curve is precisely analogous to the psychometric curve except that the underlying *state of the world* represents attributes of available choice options rather than a standardized percept such as weight. Figure 3.1(b) illustrates a case in which the placement of a given choice option on the left or right side per se makes no difference. Hence when options are indifferent, they are equally likely to be chosen, and the probability of picking the better option is the same regardless of whether it is on the left or right side. The shape of the curve now reflects the extent to which those who value one option more than the others are successful in identifying and choosing it.

3.7 Ranking Presentations by Clarity

With our measurement device in place, how can we compare the clarity of two methods of presenting the same options? What does it mean to say that one method of presentation is better than another in helping consumers make good decisions? The key idea is that clarity of presentation impacts the *steepness* of the cognitive economic curve. For a completely clear presentation device, the curve would jump vertically from zero to 1 precisely at the point of indifference. All would make correct choices even with minimal incentive. A flatter curve indicates that people have difficulties extracting the key information from the decision-making viewpoint, particularly those with relatively little at stake. Worst of all is a horizontal curve in which choice is completely random regardless of the incentive. This indicates that people find it essentially impossible to extract key information given the way options are presented.

Figure 3.2 illustrates how cognitive economic curves reflect clarity. The figure has two different curves, the steeper one being labeled clear, and the flatter one unclear. To illustrate what this means, consider the marked point $d > 0$ on the horizontal axis. This represents a case in which the left hand option is better by this amount than the right hand option. The point $P^C(d)$ on the vertical axis of the clear cognitive economic curve indicates the probability that the higher utility option is chosen with the clear presentation. The lower point $P^U(d)$ on the vertical axis of the unclear cognitive economic curve indicates the probability that the higher utility option is chosen with the unclear presentation. The relative lack of clarity of the unclear presentation is revealed by the fact that the better option is less likely to be identified. This essentially defines what it means for one presentation to be less effective in conveying important information to consumers than another.

There is in fact a great deal more that can be inferred in terms of consumer welfare from these figures, on which more to come as the

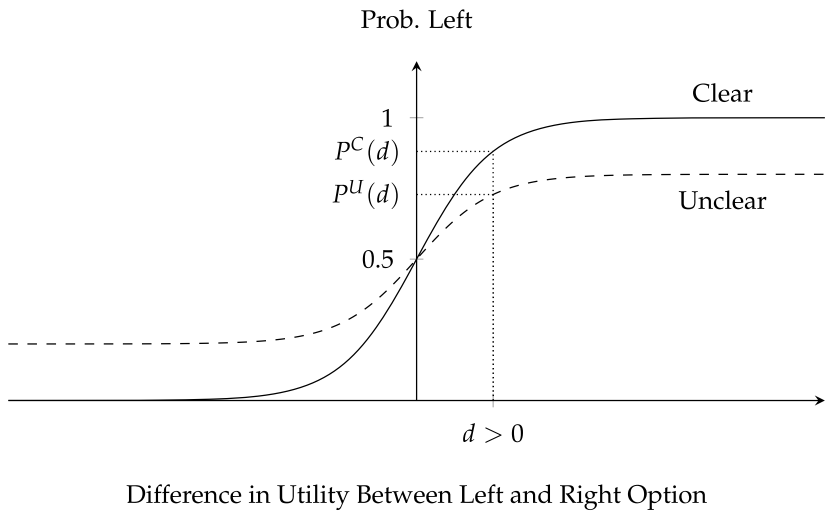

Fig. 3.2 Cognitive Economic Curves Ranked by Clarity

research moves forward. Obviously, this is just the first step of many in getting a realistic vision of the value of different modes of presentation, but it is a *sine qua non*. If we can't work forward in ideal cases, we will be unable to do so in the more intricate settings that the real world presents us with.

3.8 Online Privacy Disclosures

I close this chapter by outlining ongoing research applying new cognitive economic methods to test the clarity of *privacy disclosures*. These disclosures attempt to distill the relevant information from privacy policies to help the consumer make a decision (accept/reject conditions, or use/do not use service). Our research focuses on how the designers of these disclosures can scientifically test the efficacy of their messaging, and how regulators can step in where necessary to simplify messaging. One question we address is how much an easy-to-implement method of *indexing topics* can improve decision quality, as in the case of Mexican labor arbitration courts (Caplin et al., 2024) detailed above.

The first point is to note that privacy disclosures are very complex. One aspect relates to what data they collect about you. This can include IP address; name; email address; geolocation; payment information; sensitive info (e.g. race, pregnancy, sexual orientation); browsing history; device/browser identifier; ads clicked on, products clicked on; medical diagnostic data; and biometric data. Another important aspect the concerns uses they make of this data. This might be first-party marketing; third-party marketing; personalization (using information about you to show you what you want, as in social media algorithms); data sharing with data brokers and other websites, etc. The policy also specifies whether consenting to any of the above is required to use the website/app. In some cases consent choice has no effect at all on user experience. In other cases it indirectly affects user experience (e.g. you'll start seeing ads completely unrelated to you, which you may find bothersome/uninteresting). In yet other cases consent choice directly affects user experience (e.g. you cannot use the website without consenting, or you must pay a subscription to use the website without consenting).

Given the many details to be disclosed, privacy disclosures are far too onerous to be read in any but the most superficial manner. Even back in 2008. McDonald and Cranor calculated that it would have taken between 180 and 300 hours per year to read all of the privacy policies the average consumer consents to within a given year. I expect that to have risen many orders of magnitude by now. Bakos et al. (2014) found that roughly 1 in 1000 people actually scroll through online boilerplate when it is disclosed prior to agreement and sign in and that a median time of 29 seconds was spent by those who scrolled through the 2000 words or more of legalese.

The complexity of the privacy landscape and legitimate differences in consumer priorities make clarifying online contractual information all the more important. This is giving rise to understandable pressure to tighten disclosure rules and to validate that this allows more to make choices suitable given their priorities. Given legitimate differences in priorities, there is no universally "correct" choice. The gap that cognitive economic research is designed to fill is ensuring that choices reflect decision-makers' priorities to the greatest extent possible.

Our ongoing research begins with the simplest possible case. The starting point involves two different privacy policies and their current disclosure practices. As scientific researchers, we then identify precise contractual differences (just as psychometricians measure weight precisely). To see how well these differences are understood by consumers, we first develop an experimental *wind tunnel*. We use standard methods of experimental economics to *induce* subjects to have particular preferences over privacy options. With this we can precisely generate the corresponding cognitive economic curve for subjects offered a clean choice between these options as currently presented in the field.

How might we think about improving the clarity of the current methods of presentation? One method is to present the contract clauses in decreasing order of importance, as determined by each subject's preferences. What we would expect in this case is for the cognitive economic curve to correspondingly steepen, reflecting greater clarity. Ordering clauses in terms of importance is likely to make it easier for people to make the correct choice, hence to a steeper cognitive economic curve.

In practice, it may not be possible to order contract clauses in terms of consumer priorities, particularly as these are likely to differ across consumers. Just as in the Mexican labor arbitration cases, we may have no power to change the order of the text. Again as in that case, it may nevertheless be possible to improve presentation of options by the addition of headers that allow consumers to *self-direct* the order in which they acquire information. We will, therefore, test methods in which disclosures are not re-ordered per se, but rather clearly labeled with salient headers. We will add simple hyperlinks that can be clicked for full text and that can effectively take the role of an index simplifying identification of key information. This will allow us to test whether the indexing methods of Caplin et al. (2024) have more general value.

At some point we will have to complete the move from lab to field. This, in turn, involves several distinct steps. As part of this, we are designing surveys to identify consumer privacy preferences as accurately as possible. We are also developing multi-option variants of the two alternative fixed choice tasks underlying the cognitive economic curves. Further development of these cognitive economic research methods will be accessible using the QR code in the last chapter of the book. Stay tuned.

The methods we are designing to test for effective communication may be of value not only to regulators, but also to firms looking to scientifically test the efficacy of their messaging. It will likely be hard for regulators to keep up with changes as firms offer ever richer options to consumers. Some businesses will wish to develop reputations for clear presentation. Ideally, competition to produce greater clarity will to some extent move incentives in favor of comprehensibility. More than likely, other forces will operate in favor of opacity. Regulators surely have a role to play. So too might private businesses with expertise in cognitive economics. I sketch a corresponding market opportunity in Chapter 9.

References

Almog, D., Gauriot, R., Page, L., & Martin, D. (2024). *AI oversight and human mistakes: Evidence from centre court*. arXiv preprint arXiv:2401.16754.

Archsmith, J. E., Heyes, A., Neidell, M. J., & Sampat, B. N. (2021). *The dynamics of inattention in the (Baseball) field* (National Bureau of Economic Research Working Paper 28922).

Bakos, Y., Marotta-Wurgler, F., & Trossen, D. R. (2014). Does anyone read the fine print? Consumer attention to standard-form contracts. *The Journal of Legal Studies, 43*(1), 1–35.

Bartoš, V., Bauer, M., Chytilová, J., & Matějka, F. (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review, 106*(6), 1437–1475.

Ben-Shahar, O., & Schneider, C. E. (2017). The failure of mandated disclosure. *Russian Journal of Economics and Law, 4*(44), 146–169.

Benoliel, U., & Becher, S. I. (2019). The duty to read the unreadable. *Boston College Law Review, 60*, 2255.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review, 94*(4), 991–1013.

Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook: The cognitive domain*. David McKay.

Cameron, C. M., & Kornhauser, L. A. (2006). Appeals mechanisms, litigant selection, and the structure of judicial hierarchies. *Institutional Games and the US Supreme Court, 173*, 204.

Caplin, A., Gomberg, A., & Sadka, J. (2024). *Judging the judges: Indexing complex information reduces injustice* (NBER Working Paper).

Chemin, M., Chen, D. L., Di Maro, V., Kimalu, P. K., Mokaya, M., & Ramos-Maqueda, M. (2023). *Data science for justice: The short-term effects of a randomized judicial reform in Kenya*. Toulouse School of Economics.

Hanna, R., Mullainathan, S., & Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics, 129*(3), 1311–1353.

Oprea, R. (2020). What makes a rule complex? *American Economic Review, 110*(12), 3913–3951.

Puri, I. (2018). *Preference for simplicity*. Available at SSRN 3253494.

Shavell, S. (1995). The appeals process as a means of error correction. *The Journal of Legal Studies, 24*(2), 379–426.

Stephenson, M. C. (2010). Information acquisition and institutional design. *Harvard Law Review, 124*, 1422.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

CHAPTER 4

Cognitive Economics at Work

4.1 Cognitive Economics at Work

4.1.1 The Economic Model of Earnings

How much we earn at work is the single biggest determinant of financial well-being for most of us. That is why labor economics is such a large and important branch of economics. In recent years, a great deal of effort has been dedicated to understanding how technological change will impact earnings, with particular focus on who will be hurt and who helped by the rise of robots and computers. In response, economists are now generating ever richer information on the precise *tasks* that are involved in various jobs with a view to understanding changes in the demand for labor of different types. David Autor and Daron Acemoglu are leading the charge in this regard (Acemoglu & Autor 2011; Autor et al., 2003).

It is surely important to dedicate time and attention to understanding the precise tasks that are involved in various jobs. Yet it is equally important to understand the kinds of *cognitive skills* that are of value in the workplace of today, and, perhaps more importantly, those that will be needed in the workplace of the future. This is the subject of emerging cognitive economic approaches that I introduce in this chapter and in Chapter 6.

It is important to understand the origins of the traditional techno-centric research methods in labor economics before introducing human-centric research on cognitive skills. Economists have been studying labor supply and wages from the earliest days of the discipline. The standard

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_4

approach is to view labor productivity and wages through the lens of classical production theory. Labor is treated as a *factor of production*, along with land and capital. A production function summarizes the outputs that can be produced by efficiently combining these inputs. From this we define cost functions, which show how the prices of inputs impact methods of production. Combining this with quantitative measures of factor availability and information on demand for final products, one arrives at an understanding of the wages of labor and product prices in market equilibrium.

What matters in classical models of production is how much output can be produced with the inputs that are used. Economists avoid going onto the factory floor to look at all details, focusing instead on inputs and outputs. Those in operations management focus on factory layout, which precise machines are needed, which workers are assigned which machines, etc. But for the most part, economists abstract away from these issues to provide high level insights that apply to many contexts and regardless of the precise technology. There are inputs, there are outputs, and there is a black box production function in the middle that transforms the former to the latter.

How do economists make any progress in such an abstract framework? To some extent this progress rests on a core assumption: that of efficiency. A basic assumption in the economic model of production is that inputs are put to best use. By definition, a production function summarizes what one can produce by best practice methods of combining all inputs.

What is striking about the production function approach is how minimal it is. It goes no further into details than is required to set supply equal to demand in markets for inputs and for final goods. That is its beauty. It is also its curse. It says essentially nothing about what forms of labor are of particular value. In fact it is remarkably divorced from any human factors that might be important at work.

4.1.2 Human Capital and Cognitive Factors of Production

The first obvious limit of treating labor as a simple physical input is that it has nothing to say about why one worker might be more valuable than another. That there are differences is revealed in the market by huge variation in earnings. Digging deeply into the source of these differences is not the focus of the standard model. Hence, for example, the simplest version of the standard model has nothing to say on the face of it about whether or not it is worth paying for an education.

It is the observation that it makes no obvious allowance for education and its impact on wages that led to a brilliant amendment to the production function approach to earnings. Becker (1964) made a far-reaching analogy between physical capital and human capital. Just as one needs to invest resources up front to build up physical capital, he pointed out that one needs to sacrifice wages and pay for an education to build human capital. The decision on whether or not to do this depends, as with any investment, on the balance between cost and benefit. Even questions such as what to learn can be captured by models that balance differential rewards against the corresponding costs.

How can one apply the theory of human capital to explain real labor market phenomena? In principle, this can be accomplished in standard economic stories. The best documented change in recent decades has been the growing premium commanded by those with a college education even while the supply of such workers grew (there is some evidence that the tide is turning in this regard, but only time will tell). The theory of human capital allows us to rationalize changes in the wage premium over time just as we would that of any other input to production: as resulting from shifts in the technology of production and the availability of complementary factors, such as physical capital.

Valuable as it might be to interpret history through this lens, it does not link in any obvious way to durable skills that will be of value as technology changes. Hence it has little to say about what humans should be good at to thrive in the labor markets of the future. This will surely become ever more important as many of our decision-making responsibilities are taken over by artificially intelligent machines. We need to pick up the ideas of Welch (1970) and Jovanovic and Nyarko (1995), and model *cognitive factors of production*.

An early hint of the importance of cognitive factors in earnings is provided by Mani et al. (2013), who argue that poverty affects work ability and earnings capacity through a variety of channels that deplete attentional resources and negatively impact job performance. In support, Kaur et al. (2021) pay poor subjects a piece rate for completing a repetitive yet intricate attention-demanding task. They find that workers who are paid earlier are better able to complete these tasks and consequently earn higher incomes. Other poverty-related factors that have been found to affect attentional constraints and income are sleep deprivation, exposure to high levels of air pollution, exposure to high levels of noise pollution, and unsolicited workplace interruptions. What is not clear is how to go beyond case studies to study human elements more broadly.

Drilling down into the cognitive factors that contribute to earnings over the course of an individual worker's career is the central charge of the cognitive economic approach. From this perspective, human capital theory is an ingenious method of ducking important questions. It deliberately side-steps the issue of precisely what education produces that is of value to employers. This has become harder to sustain in recent years, and looks set to get even more challenging in the years to come. Ongoing technological change is threatening to entirely change the skills requirements of jobs. The limits of the standard approach are coming into sharper focus as much of what we do is being replaced by machines such as robots that are in many respects dominant. There is increasing interest in understanding precisely why humans might remain essential to production processes before we are all replaced.

4.1.3 Decision-Making Skills: The Measurement Challenge

In an important conceptual innovation, Deming (2021) makes the case that what will matter to earnings going forward are *decision-making skills* that are generalizable across tasks, rather than skills at specific tasks. These *higher order skills* are by definition not attached to any given task in current production processes but rather allow better adaptation to changing task needs. Deming links them explicitly with the higher levels beyond stage 3 in Bloom's taxonomy of educational objectives: those of analysis, synthesis, and checking against changing external realities.

As so often in cognitive economics, the measurement challenges are profound. To understand the nature of these challenges, one needs to understand the methods of measurement that define modern labor economics. Central to the field is the now essentially limitless, in best cases close to complete, data on patterns in earnings and employment over the life cycle. These are used to identify key patterns, such as increases in inequality, returns to schooling, the impact of layoffs, etc. Our knowledge of these job histories has grown massively in recent decades with booming availability of corresponding data. Many surprises are being digested about these histories as part of the essential groundwork for understanding what factors determine labor market success.

Yet even complete data on labor market outcomes provide little insight into cause and effect. Economists have methods to overcome this. One method that is central to modern labor economics involves event studies and/or randomized control trials. With regard to the former, labor economists have become experts at identifying "natural experiments". By definition, these satisfy formal statistical criteria that allow separation of cause and effect. Yet by their nature, there are limits to what these studies can teach us about individual differences. It is typical in such studies to focus more on the jobs and the industry in question than on what happens over time to the workers. For example, one can identify how an increase in the minimum wage impacts employment and earnings *in an industry* without any need or ability to dig into differences in outcomes *at the worker level* and how they play out in the response to disturbances that all for flexibility of response. This is well understood in the literature, which identifies *average* treatment effects rather than *individual* treatment effects as would be required for identifying the importance of individual differences in decision-making skills.

How can one go beyond these standard approaches in labor economics to an understanding of the role of decision-making skills? To understand the depth of the challenge it is worth contemplating what we would ideally like to measure. This ideal is not only incredibly rich but also profoundly intrusive. To identify the role of decision-making skills in lifetime earnings, we would first need to conceptualize their role clearly, and then identify the trace of these concepts in and out of the workplace. Effectively we would have to remove any constraint of privacy. Not only would we need to follow workers around in their workplace, but also while working from home and in their virtual interactions. We would also need to track their activities outside the workplace. We would want to capture their exploration of outside options, their pursuit of opportunities to increase training, considerations related to time out of work, discussions of impending job threats, etc.

I believe that we would have much to learn from this type of detailed ethnographic fieldwork. Yet it is hard for me to imagine exactly how it could be organized, who would be able to conduct it, and how researcher beliefs and biases might interfere with scientific investigation. For the moment at least, we are left to explore alternative approaches.

Given the challenges, first evidence on the role of higher order skills derives in large part from data on job postings and on task-based job decomposition. Those looking for evidence that higher order skills can identify corresponding patterns in job advertisements, which increasingly ask for such skills rather, somewhat to the expense of standard educational credentials (Deming, 2021). This pattern suggests that employers want workers who will flexibly respond and adapt to unforeseen circumstances, such as changes in job requirements. A related advance in measurement derives from the increasing focus on identifying the tasks associated with different jobs. As part of this, labor economists have created measures of the *decision-making intensity* of jobs. The evidence shows that, in broad terms, jobs that require higher levels of decision-making skill are better paid and growing faster than those at the lower end of the scale. But again, these forms of evidence speak to job requirements not the differential abilities of humans to fulfill these requirements and where they stem from. An altogether different approach is needed focused on these human elements.

4.1.4 A Research Template for Cognitive Labor Economics

The goal of cognitive labor economics is to model and measure the impact of various forms of decision-making skill on labor market outcomes. For all the ingenuity that has been applied to enriching standard labor market data, it is hard to make quantitative progress in understanding which particular skills are important for earnings without to some extent measuring these skills at the individual level. Without that, while there is much *indirect* evidence that focus on the role of higher level skills is appropriate, there is little by the way of direct evidence. In his recent review of human capital theory, Deming (2022) summarizes roughly where we stand and what the challenge is. While it is qualitatively clear that higher order skills, such as problem-solving and teamwork, are increasingly valuable, we have not done a good job either in measuring these skills, or in working out their importance for patterns in life-cycle earnings and employment. This means further that we know very little about how these skills develop and the extent to which they can be taught. To change this we need somehow to identify key individual differences in decision-making skills.

What is clear at this stage is that progress requires new modeling and measurement paradigms drawing both on the psychological and economic traditions. In what follows I outline settings in which this kind of work is under way in earnest. I focus in particular on design and fielding of instruments that elicit *individual differences* in key decision-making skills and how they relate to labor market histories. The new forms of measurement that are being introduced draw strongly on the psychological tradition in personality measurement. Particularly well-developed personality inventories relate to the "Big Five "personality traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism. There have been many efforts to tie these measures with economic outcomes, with mixed results. There are also some general correlations between earnings and tests of intelligence, such as recognition of patterns in numerical sequences. While these are broadly predictive, they do little to identify particular areas of decision-making in which this form of pattern recognition is helpful. They are also divorced from decision-making per se.

Ongoing research links tests of the ability to take in complex information with tests of how well that information is deployed in decisions. As in Bloom's taxonomy, the minimal level of understanding that is of interest is not factual, but rather the translation of facts to effective decision-making. This defines the key structure of the cognitive economic approach to eliciting decision-making skills that are of particular value for life-cycle patterns of earnings and employment. In its ideal form it is a 5 stage process:

- *Stage 1. Model Formulation:* This involves conceptualizing and modeling a cognitive skill to isolate and study. This requires a major conceptual innovation. It switches focus away from a standard production function approach to labor productivity to an *information-theoretic* approach.
- *Stage 2. Experimental Implementation:* This stage designs and experimentally implements instruments to assess individual differences in these skills and measures first order correlations with patterns of earnings and employment.
- *Stage 3. Administrative Implementation:* This stage implements instruments in contexts in which appropriate administrative data is available.

- *Stage 4. Training Protocols:* This stage develops and tests instruments to teach the identified skills.
- *Stage 5. Implementation in the Field:* This stage implements teaching protocols in the field to identify how they impact the actual evolution of earnings and employment.

The first implementation of this staged process is ongoing in research on economic decision-making skill, as now outlined, we are advancing to stage 3 in this first case, based on successful completion of the first two stages.

4.1.5 Economic Decision-Making Skill and Comparative Advantage

Caplin et al. (2023) implement the cognitive economic research process for a particularly central economic decision-making skill. The starting point for this research is a conceptual specification of an abstract decision-making skill that should have an impact on job performance, wages, and transition to management roles. What might that be, and why is it so important?

In answering this question, we took our lead from Paul Samuelson. When he was asked to name a central but non-obvious implication of economic theory, he pointed to the theory of *comparative advantage*. We follow that in defining our measure of decision-making skill, which requires both raw information processing capacity and an ability to use information strategically by exploiting comparative advantage.

What is the technical translation of this idea into a decision-making task? We model the amount of output produced by a fixed set of inputs depending on how well the decision-maker understands the true nature of productive opportunities. Making good decisions involves making trade-offs. Which worker should a manager assign to which task? Which questions require immediate attention and which can be delayed? The model features a decision-maker strategically acquiring information about factor productivities under time and effort constraints. To make the goal of measurement precise, we specify a setting where agents assign factors of production to different roles to maximize total output. This could be a manager assigning workers to jobs, or workers allocating their own effort to job tasks. Factors have heterogeneous productivity schedules, so the agent must compare hypothetical assignments and choose the one with the highest expected output. Agents acquire costly information about payoffs to different actions, deploying their attention according to their idiosyncratic level of skill.

Our research develops a theory and measurement paradigm for assessing individual variation in economic *decision-making skill*, which we define as the ability to make good decisions about resource allocation. We see this as a central skill for being productive at work. Viewed abstractly, many work-related decisions are allocative and require appropriate decision-making based on an accurate assessment of options and their impact. For example, one's work day is a fixed resource in need of being allocated properly. One must allocate one's time and effort not in line with absolute skills, but rather by comparison of available alternatives and relative skills. Our thesis is that those with high such skills will have success in the labor market in a manner that will carry the signature of good decision-making, such as higher income and more decision-making responsibilities. Modeling decision-making skill requires us to treat attention as a scarce resource that some people deploy more effectively than others. As a consequence, we must use the tools of information theory to measure individual differences in labor productivity.

Technically, the way we define individual-specific attention costs is strongly analogous to the role of input costs in standard production theory. In a competitive labor market, workers with higher earnings per unit of time (e.g. wages) have a higher marginal product of labor. In our model, agents with higher levels of economic decision-making skill choose more efficient allocations, holding constant information complexity and time constraints. Decision-making skill captures total processing bandwidth, but also the ability to strategically pay attention to important information and to understand comparative advantage. Our research contributes to human capital theory by formalizing the value of this skill in the labor market. To continue the analogy with production theory, what high skills produce is better decision-making. This may well be a key skill for the cognitive era. Going forward our research will dig into many different forms of decision-making skill and their link with earnings over the course of a career.

We measure economic decision-making skill in experimental settings by creating a novel task we call the Assignment Game. Participants are managers who assign fictional workers to jobs to maximize output. They observe multiple draws from workers' productivity schedules over tasks and then make an assignment. Participants are scored based on each worker's mean output in the task to which they were assigned. The Assignment Game requires participants to process information quickly and to assign workers to their highest value task given the skills of the others.

There are a few points to note. First, the actual payoffs to each possible assignment are precisely defined. This means that we can cleanly identify errors using methods introduced in the last section that draw on the psychometric tradition. The distinction is that rather than using these data to measure differences in the quality of presentation devices as there, in the case of allocative skills we use them instead to identify individual differences, which turn out to be substantial. As the experimental designers, we know the payoff to all possible allocations. We then see the choices that subjects make. Their failure to pick optimally is visible in each decision problem that they face. In essence, it is the extent of this failure that reflects what we measure as their individual-specific decision-making skill. On average, those with high such skills make choices that are closer to optimal than do those with low such skills.

In essence, we see our Assignment Game as an economic IQ test. Typical IQ tests are essentially logical puzzles with no link to any economically interesting decision. By contrast, our design is a quantitative puzzle that requires an understanding of the principle of comparative advantage, possibly the most central idea in the theory of production. One must assign not according to absolute productivity of a factor in a task, but rather their ability in that task relative to their and others' abilities in all tasks. It is hard to think of a more basic economic problem than this. You are welcome to try.

4.1.6 Economic Decision-Making Skill and Earnings

We have implemented our experimental paradigm both in the U.S. and in Denmark. The precise instruments are somewhat different given the very different pool of respondents. In the U.S. respondents are paid participants in the Prolific platform, a popular platform for conducting online experiments. Panel members are more than happy to spend up to an hour on our interface, given that they are incentivized and earn side pay for their time. In stark contrast, our Danish sample is entirely voluntary, and our instrument has to be significantly shorter. What is striking despite the very different settings in samples is that the main findings are powerfully aligned, as now outlined.

Taking first the U.S. case, we have a sample of more than one thousand full-time U.S. workers ages 25–55. In addition to the Assignment Game, participants also completed a demographic and labor market survey, which elicited information about current income and occupation. Our findings reveal significant individual differences in economic decision-making skills that correlate in intriguing manners with patterns in income and employment.

The key research questions concerns dynamic linkages between economic decision-making skill and patterns of earnings and job transitions in administrative data. The first research question is how differences in this skill relate to current earnings. Early results are encouraging in this regard. We find that decision-making skill is strongly associated with income, even after controlling for IQ, numeracy, education, occupation, and other covariates.

Perhaps most strikingly, we also find that the association between decision-making skill and income is significantly greater in decision-intensive occupations. These are defined using the task-based methods of decomposing jobs of Autor et al. (2003) and Acemoglu and Autor (2011). Since the value of decision-making skills is grounded in economic theory and strongly predicts economic success even conditional on IQ and other measures, the Assignment Game can justifiably be viewed as an economic IQ test.

A limitation of the U.S. study is that we know little about the earnings histories of respondents. What we do know derives from survey questions that we posed while gathering responses. It is obviously important to get more granular information. Luckily we are well-positioned to do just that. Encouraged by first experimental results, we placed simpler versions of the Assignment game instrument into a setting that permits us also to measure life-cycle patterns of income. In fact this next stage of research on decision-making skills is being conducted in the most important modern sources of panel data in labor economics: the Scandinavian population registries. The Danish registry infrastructure is particularly suitable since it allows us to direct survey instruments appropriately, to know relevant details of respondents' histories and those of any businesses they manage, and to track their future outcomes (see Andersen & Leth-Petersen, 2021; Epper et al., 2020). As indicated above, our research design faced novel design challenges given that we could not offer rewards so that completion must be strictly voluntary, so that the survey instrument had to be simplified substantially.

First results from Denmark are now in, and the findings are strikingly similar to those in the U.S. Measured decision-making skill predicts earnings, with the link being stronger in decision-intensive industries. What makes the results particularly striking is that we left ourselves no degrees of freedom in Denmark. We applied precisely the same econometric methods to estimation in Denmark as in the U.S.

Given the positive and provocative first findings, the research team is now putting in place enough measurement devices to learn much about the evolution of allocative skills over the course of a career. We have measured these skills together with many other forms of intelligence among a sample of Danish twenty year olds. We have done the same for a group of retirees. The registries allow us to link these with past and future labor market outcomes, and even to repeat tests and run other interventions in later periods. Stay tuned for updates.

4.1.7 A Procedural Amendment

It is worth commenting on what research teaches that was not anticipated at the design stage. For purposes of clarity, in our experimental design we presented information on worker productivity sequentially. We first presented information on how good worker 1 was at all tasks. We then did the same for worker 2, worker 3, and so on.

While our first skill measures are based on a simple rational model of costly learning that captures the difficulties subjects have absorbing rich information regardless of order, there are plausible decision-making short cuts that make order of presentation important. Our experiment identifies just such effects. Many, particularly those who are not good at the task, appear first to place worker 1 in their most productive position, then place worker 2 in their most productive unfilled position, and so on. There is significant variation both in how often different subjects use an algorithm like this, and how well they do overall. Strikingly, for those who use the sequential procedure a great deal, the relationship between measured skill and earnings is relatively weak. They appear more to have been lucky when they used the procedure rather than revealing their skill in deciding which assignments are better overall.

Given this finding, our implementation of the decision-making skills instrument in the Danish registry has been designed to provide insight not only into the quality of the final decision, but also the nature of the decision-making process. We are developing a model that mixes cognitively sophisticated methods and heuristic methods to understand patterns in the resulting data, in the spirit of Arrieta and Nielsen (2023). The million dollar question is how the differential skills that our instrument will allow us to estimate will play out in data on patterns of earnings and employment over the life cycle. Your guess is as good as mine. That is research for you!

4.1.8 Teamwork, Social Skills, and Earnings

There is much justifiable interest in the part that social and teamwork skills play in earnings. The measurement challenge here is that workers do not get randomly assigned to teams with team outputs being monitored as measures of individual contributions would require. This is the background to the study of Weidmann and Deming (2021) which generates experimental data precisely to overcome this constraint. Their key innovation is to translate social skills into output improvements by conceptualizing their impact on work output of teams. In this setting the nature of social skill is to facilitate improvements in team productivity over and above the contribution based on individualistic skills alone. Team players are then thought of as individuals whose addition to a team does most to boost the productivity of other team members.

Weidmann and Deming design tasks that involve cognitive challenges that can be administered with only minor modifications both to individuals and to groups. Tasks have clear correct answers so that they can definitively measure performance of individuals and groups. This enables them to estimate group performance controlling for individual task-specific skills. Given their interest in team-play, these are tasks in which it is reasonable to expect cooperation among group members to improve team performance. The key finding is that there are individuals whose contribution to group cognition is significantly higher than their direct cognitive skills would suggest.

There is a fascinating link with the work of Silver, Mellers, and Tetlock (2021) who undertake a careful laboratory study of how communication impacts the quality of decision-making. All subjects answer a fixed number of difficult quantitative general knowledge questions (e.g. the population of Uzbekistan). The ensuing experiment has three phases. In the first stage, they are left to answer the questions alone, with incentives. In addition to providing answers, they are asked to rank their levels of confidence. In the second stage of the experiment, subjects gather together with their pre-assigned group and communicate freely. In the final stage, they each return to their workstations and enter a final answer and rank their confidence in that answer.

The results reveal a split. Some group meetings result in improved performance, with guesses not only moving closer to one another but closer to the truth on average. In other cases, while guesses move closer to one another, they end up on average further from the truth. Analysis of the confidence data produces a provocative and interesting hypothesis about learning. The hypothesis is that the more confident individuals generally have a higher influence on the learning that takes place in the group. What this means is that the actual skill level of the confident individuals matters. What makes a group communicate effectively and learn is that those who are confident are so appropriate. What makes a group communicate in a manner that degrades performance is that those who are confident are comparatively bad. Their false confidence appears to be persuasive, a theme that will play out intriguingly in relation to crowd-sourced learning in the next chapter.

Measuring task performance before and after team communication suggests the hypothesis that team players aid the communication process in some manner, perhaps by ensuring that overconfident low ability loud mouths do not monopolize communication, or perhaps by helping the group better differentiate between more and less (objectively) informative communication. Note that the results of Silver, Mellers, and Tetlock suggest that it may be important to identify differences in confidence and overconfidence in relation to teamwork skills. This is a subject that will be revisited at some depth in Chapter 6 in relation to skills working not only with other humans, but with AIs in the workplace of the future.

There are important and under-explored questions relating to how differences not only in skills but in attitudes impact team-play. The literature on *social preferences* surveyed by Fehr and Charness (2024) highlights the impact of moral norms on team-play, in particular the extent to which individuals are willing to free ride on others efforts. In the case of teamwork in particular, there needs to be convergence between research on decision-making skills and on moral norms. No one likes to feel taken advantage of. The presence of free riders may effectively lower the degree to which other team members are themselves willing to contribute.

References

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043–1171). Elsevier.

Andersen, H. Y., & Leth-Petersen, S. (2021). Housing wealth or collateral: How home value shocks drive home equity extraction and spending. *Journal of the European Economic Association, 19*(1), 403–440.

Arrieta, G., & Nielsen, K. (2023). *Procedural decision-making in the face of complexity*.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics, 118*(4), 1279–1333.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Caplin, A., Deming, D.J. Leth-Petersen, S., & Weidmann, B. (2023). Allocative skill. National Bureau of Economic Research. NBER Working Paper 31674.

Deming, D. J. (2021). The growing importance of decision-making on the job. National Bureau of Economic Research. Working Paper 28733.

Deming, D. J. (2022). Four facts about human capital. *Journal of Economic Perspectives, 36*(3), 75–102.

Epper, T., Fehr, E., Fehr-Duda, H., Kreiner, C. T., Lassen, D. D., Leth-Petersen, S., & Rasmussen, G. N. (2020). Time discounting and wealth inequality. *American Economic Review, 110*(4), 1177–1205.

Fehr, E., & Charness, G. (2024). *Social preferences: Fundamental characteristics and economic consequences*.

Jovanovic, B., & Nyarko, Y. (1995). A Bayesian learning model fitted to a variety of empirical learning curves. *Brookings Papers on Economic Activity. Microeconomics*, 247–305.

Kaur, S., Mullainathan, S., Oh, S., & Schilbach, F. (2021). Do financial concerns make workers less productive? National Bureau of Economic Research. Working Paper 28338.

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science, 341*(6149), 976–980.

Silver, I., Mellers, B. A., & Tetlock, P. E. (2021). Wise teamwork: Collective confidence calibration predicts the effectiveness of group discussion. *Journal of Experimental Social Psychology, 96*, 104157.

Weidmann, B., & Deming, D. J. (2021). Team players: How social skills improve team performance. *Econometrica, 89*(6), 2637–2657.

Welch, F. (1970). Education in production. *Journal of Political Economy, 78*(1), 35–59.

CHAPTER 5

Cognitive Capital and Human-AI Interactions

5.1 The Human-AI Decision-Making Pipeline

The upcoming cognitive revolution, like its industrial precursor, is defined by a radical change in technology that is uprooting traditional social structures and forms of work. Yet social science still bears the imprint of that earlier revolution, with its focus on physical factors of production, such as labor and capital. The machines of today introduce intangible cognitive capital that substitutes for human mental effort. In the next few chapters, I illustrate how cognitive economic methods can a help us understand the impact of the cognitive transition and liberating the great promise that the AI revolution brings with it, while avoiding the many pitfalls.

My focus in this chapter is on designing human-AI interactions to produce better decisions. Medical diagnostics, specifically cancer-related, is the running example. Typically different courses of treatment are appropriate depending on the diagnosis. Options may include giving a clean bill of health if cancer can be definitively ruled out, operating aggressively if the cancer is at an advanced but treatable stage, conducting chemotherapy, etc. What AI's offer is the ability to interpret massive volumes of data to look for patterns that predict the current diagnosis and its likely evolution under different treatment options. AIs can process far more information than can any one diagnostician, and in the process extract patterns that are not readily visible even to experts.

While there are those who foresee AIs one day taking over all such diagnostic decisions from humans, I outline in this section why that day

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_5

is far off and indeed may never arrive. To understand what the main challenges are it is useful to separate out three stages of human interaction in diagnostic decisions in a *human-AI decision-making pipeline*.

1. *Data Generation:* Ideally, machine learning models generalize from cases that have been pre-classified: e.g. cancerous or non-cancerous for a skin lesion. These are the *ground truth labels* on which all subsequent inference is based. Unfortunately in the vast majority of applications, medical included, there are far too few cases for which definitive labels are available. It is something of a dirty secret that even the best-known datasets used in competitions are riddled with mistakes. In an article revealingly entitled "Everyone wants to do the model work not the data work: Data Cascades in High-Stakes AI", Sambasavian et al. (2021) present case studies in medical and other important applications that highlight the profound damage that results from low-quality data work. Ongoing research outlined in Sect. 5.3 shows that cognitive economic methods can do a great deal to *improve data quality* and address this challenge.
2. *Aligning AI Values with Human Values:* Once there is a labeled dataset, the data scientists tell everyone else to leave the room while they pursue their dark arts. They end up with models that are typically far better than the average diagnostician, so do cognitive economists like me have to raise issues? In Sect. 5.4, I outline recent research showing that there is much that cognitive economics can contribute to improving model performance. The key observation is that there is a profound *alignment problem* in getting the AI to reflect human priorities. This is related to, yet distinct from, the current conception of this problem in machine learning Christian (2021). The key challenge analyzed in this literature is that of getting the AI to learn human preferences. The application of cognitive economics shows this problem to be *bilateral*. It is also important for the data scientist to understand the motivations and constraints of the AI. I also outline cognitive economic methods of resolving this problem based on *revealed preference* reasoning.
3. *Action Selection:* There are many joint architectures for combining human and AI information in arriving at actionable diagnostic decisions. In typical medical cases, it is standard for a skilled diagnostician to have the last word after having been provided with

algorithmic summaries of much of the relevant evidence. For this to work well requires human experts to understand what the algorithms are doing well enough to use them effectively. What little evidence there is on how this works in practice is discouraging with evidence of widespread *algorithmic aversion*, which is believed to flare up when obvious mistakes are made by the algorithm that no human would make causing global dismissal even of good algorithmic advice Tejeda et al. (2022). In Chapter 6, I outline research on the skills that allow people to work effectively with AI.

Note that the research outlined in the main body of the book treats each stage in the decision-making pipeline separately. In Chapter 9, I propose holistic research integrating all stages. This is vital to avoid a worst case cascade of biases in which low-quality labels, poorly chosen loss functions, and inappropriate human-AI interactions ending up producing poor quality and biased final decisions that are even more exaggerated than those made without AI assistance. A current day sketch of a human-AI decision-making pipeline would reveal a Rube Goldberg kluge. We work silo-by-silo without any clear view of their final impact on decisions. Cognitive economic research treats all stages in a coherent and unified manner.

5.2 Human-AI Interactions AI and Cognitive Economics

There are three fundamental reasons that cognitive economics has so much to contribute to the design of human-AI interactions. The first is that the economic model of learning is of huge value in organizing understanding of the decision-making pipeline. The second is that the dataset that is standard in diagnosing performance of AIs is essentially ideal for cognitive economics. The third and final reason is that one can readily implement cognitive economic methods in lab experiments to test out design principles to improve the quality of final decisions. I take these up in turn.

Let me first address the modeling issue. As you might expect, economists model learning as a black box input-output device. The input is all of the relevant information. The black box can be a computer, a human, a crowd of humans, all in tandem, etc., that transform this unfiltered information to the point that it is actionable. The output of this black box is the final choice of action, presumably better informed than purely random choice. Applying this to the medical setting, a new case of possible cancer is considered. Various forms of learning are applied to the evidence. The case is then classified as the learning suggests and the treatment decision made. The goal of the information processing stage is to learn enough to make decisions that are appropriate to the diverse cases being considered.

Given that learning is imperfect, mistakes are sure to be made. The economic model specifies the role of all learning, by human, machine, or any combination of these, as being to reduce damaging mistakes. In the case of cancer, one needs to know how bad it is to give a clean bill of health to a patient who in fact has early or late stage cancer, and on the converse side how bad it is to send a healthy patient on a long wild goose chase through the thicket of medical tests due to a false positive. The optimal method of combining human and AI inputs is then defined by a comparison of cost and benefits. Additional human and AI resources should be deployed until the marginal benefit no longer exceeds the marginal cost.

How does this abstract economic viewpoint help? For one, because it presents a unified picture of the human-AI decision-making pipeline. A second basic lesson the economic model conveys is that adding together diverse forms of learning intelligently must improve the quality of final decisions. It is reasonable to conjecture that we are wasting massive resources and far from the efficient frontier in terms of how much we improve decisions as a result of the investment we make in human and AI modes of learning. A third lesson that the model teaches us is that the optimal resources to apply to improved learning must balance resource costs against benefits in terms of loss reduction. Once again, because we have not seen the decision-making pipeline in a holistic manner, we do not know the incremental value of additional resources at each stage of the decision-making pipeline.

It is one thing to criticize current approaches. It is quite another thing to improve on them. Fortunately, cognitive economic methods have much to offer of practical value. The key to this is a remarkable stroke of good fortune. *The dataset that is standard in evaluating AI performance in classification tasks is also ideal for cognitive economics.* The tradition in judging algorithms is to hold out a *test set* to confirm that out of sample performance is approximately as good as in sample performance. This avoids over-fitting. Once a good enough such model has been developed it is taken into the field and used to classify novel cases based on all it learned from fitting the test cases.

In judging between AI models, the comparison is between algorithmic and ground truth classifications. The data on which performance is judged is the *confusion matrix*, which pinpoints the nature of the classification mistakes made by each model that is investigated on the path to arriving at the final model used for classification. This same form of confusion matrix can be used for judging performance at all stages of the decision-making pipeline and indeed for any complete pipeline.

What does this have to do with cognitive economics? To understand the link, think in terms of the SDSC dataset that is so central to cognitive economics. As noted in Chapter 3, this form of data is very helpful in assessing choice mistakes and how easily true underlying differences between choice options are perceived. It was introduced in that section precisely to show that one can design measurement protocols to identify which choices are and are not mistaken. In the case of an AI algorithm, we are likewise interested in mistaken classifications that are made when the algorithm mis-classifies cases, e.g. declaring an image of a cancerous cell to be non-cancerous, and vice versa. In cognitive economics and psychometrics, SDSC reveals cognitive constraints. So it does again in the case of algorithms: it reveals the AI's limited ability to identify ground truth. The fact that precisely the same dataset is used in judging human and AI cognitive limits underlies the important role that cognitive economics can play in understanding human-AI collaboration.

The third and final contribution of cognitive economics to machine learning lies in the experimental lab, which can be used to understand critical aspects of human-AI interaction in reaching categorical decisions. SDSC is frequently gathered in experimental laboratories. This means that one can generate experimental data to compare the performance of any method of classification, however, humans and AIs may interact in producing it, on precisely the same level playing field.

5.3 Cognitive Economics of Data Generation and Labeling

As noted above, to train algorithms effectively requires many more ground truth labels than are definitively available. This is essentially universal in medical decision-making and very widely so in many other applications, such as self-driving cars, development of generative AI, etc. The end result is that many of the supposedly objective labels from which algorithms generalize are in fact generated by human judgment. The need to create additional high-quality labeled data using human input has given rise to a large and rapidly growing labeling industry that pays members of their community for labeling and annotation tasks vital to self-driving cars, generative AI, and other areas built on massive amounts of labeled data.

In addition to general purpose, platforms such as Scale AI and Amazon M-Turk, there are specialized communities for medical labeling, such as Centaur Labs, which is a research partner. Many members of the Centaur community are medical students and even medical professionals looking for additional income. Ongoing interdisciplinary work with Jennifer Trueblood, Bill Holmes, Gunnar Epping, Daniel Martin, Philip Marx, as well as Erik Duhaime of Centaur labs is developing new methods of labeling based firmly on cognitive economic principles.

Our point of departure is the current standard practice in the labeling industry that uses crowd-sourced *categorical* labels, e.g. whether or not an image of a blood cell indicates the presence of cancer. Much like when voting in elections or answering multiple choice questions (more on that later), there is no room to hedge bets. Many of the sources of error in individual decisions are idiosyncratic, and these get averaged out in crowd settings. Of course, there are typically limits: if even the most expert pathologists disagree, a show of hands by less informed labelers will be of little value.

Our first scientific question in applying cognitive economic methods to the labeling industry concerns the possible value of enriching the response options for crowd members by using *elicited belief* methods. The idea is simple. Well as current approaches perform, the economic model of learning suggests that they throw away valuable information. Being certain of the truth is the exception, not the rule. When a labeler takes a quick first look at an image, they may have doubts, and there is no reason to believe that most fully resolve their uncertainty for all images.

There is a great deal of randomness in individual performance in even the simplest of tasks, such as judging which of two noises is loudest. Can one realistically expect this to be absent in judging whether an image of an irregularly shaped blood cell is or is not cancerous?

The obvious implication is that it might be of value to allow labelers to express their uncertainty in some useful manner. The most useful of all is a probabilistic expression, with a probability score of 75%, for example, indicating the belief that this cell is 75% likely to be cancerous. In support of this approach, it is standard procedure in experimental economics, and much research has been conducted on how best to elicit subjective beliefs. Key to this are methods for measuring subjective beliefs in an incentivized manner using *proper scoring rules* that reward subjects for correctly revealing their beliefs. Danz et al. (2022) summarizes state-of-the-art experimental procedures.

The open research question is whether and when the use of proper scoring rules enhances the wisdom of the crowd. On the pro side, they allow labelers more latitude in quantifying their confidence in their opinion. In that sense, beliefs are more informative in a sense made clear by David Blackwell (1953), and hence have the potential to improve decision-making. Yet this clean theoretical vision needs two heavy caveats. First, if respondents find it more difficult to report their beliefs rather than simply reporting the most likely category, some may opt out, and those who do participate may absorb different information than they would had choice been forced. Second, behavioral economists have shown that subjective probability assessments can be out of line with reality. Following the pioneering work in this regard of Kahneman and Tversky (1979), *probability weighting* has been found to be widespread (see also Benjamin [2019]). An individual exhibits extreme *overconfidence* if they are always sure about a classification but are correct only 51% of the time. They are *underconfident* if they are only 51% sure, but are always in fact correct. While both forms of miscalibration have been identified in practice, overconfidence appears to be more widespread. Only those that are *well-calibrated* have subjective probabilities that closely mirror actual probabilities: when they are 90% sure of a classification, they are right 90% of the time. If those who are the *least skilled* are *the most overconfident*, then weighting beliefs according to subjective confidence allows ignorant but confident types to overrule those who are both better informed and better calibrated. Sound familiar?

With no a priori obvious answer, the key is to conduct a field study. We have completed the first such study with the Centaur community on its regular Diagnosus app on classification of white blood cells that may or may not be cancerous. The cases under study have all been ground truth classified by unanimous agreement of three medical experts at Vanderbilt University Medical Center. The experimental subjects providing labels are members of the Centaur labeling community, many of whom are medical students. Subjects who enter our blood cell competition are told that the images they will be shown are equally as likely to be cancerous as not. After training they are randomly assigned to either the categorical choice or elicited belief condition.

The results are now in and the bottom line is simple. Gathering elicited beliefs substantially improves the quality of medical labels. A three-way decomposition of the overall effect reveals how the use of elicited beliefs improved labeling, and what might and might not be general about the particular case we study.

1. *The Categorical Learning Effect:* Being asked to pick a category as opposed to provide a probability might make it less important to resolve uncertainty one way or another. In practice in the case of the blood cells, it turns out that changing from forced choice to elicited beliefs does little to lower the quality of learning. We anticipate that in other cases categorical learning might be worse with elicited beliefs.
2. *The Confidence Effect:* The real novelty in gauging elicited beliefs lies in their quantification of subjective confidence. When elicited beliefs produce a different verdict than majority rule, it can only be because a *confident minority* overturns a *less confident majority*. On average in our experiment, this turns out to be the right thing to do since there is a positive correlation between confidence and accuracy. One might expect this to be a relatively general finding.
3. *The False Confidence Correction:* What recalibration adds to the picture is that it allows us to correct for false confidence. When recalibrated elicited beliefs produce a different verdict than raw elicited beliefs, it must be because a *justifiably confident* group of labelers on one side of the 50% line overturns the verdict of an *unjustifiably confident* group on the other side. More often than not, this is the correct thing to do in our setting, and perhaps more generally.

There are plenty of other low-hanging fruit in this area. There is a complementarity between labeling and machine learning. By way of illustration, the AI can itself generalize from probabilistic labels rather than categorical labels. To do this require re-calibrating responses *after counting votes or calculating means*. Early evidence suggests that this technique further boosts algorithmic performance.

Pulling back, the labeling industry is possibly the most pure illustration of cognitive economic principles that one can find. Design of the payment scheme for incentivizing labelers has many important roles. One would like the incentive scheme to reward attentional effort for all different skill levels. How responsive people are to rewards for getting classifications correct can be tested in simple experimental protocols. Going one step further, to choose the right crowd given a payment scheme is to assess the marginal value of individuals to crowd decisions and to select the crowd by equalizing value at the margin with the cost for each skill level. There are cost-quality trade-offs, since a crowd of experts is too expensive and cannot supply labels in appropriate volume. To identify the marginal value of each contributor requires further experimental protocols in which one varies team composition. Much exciting work lies ahead.

5.4 The Bilateral Alignment Problem

Let's move on now to the second stage of the human-AI decision-making pipeline, when data scientists grab the data file and work to find the best model to classify medical cases. In doing this they generally make intuitive efforts to reflect the underlying human values. Suppose we are interested in predicting pneumonia from chest X-rays. Due to the asymmetric health risks stemming from failing to detect a case of active pneumonia, our diagnostic procedure is very conservative. We are willing to accept 99 false positives for every one false negative. Reflecting these preferences, Caplin et al. (2022) train CheXNeXt, a 161-layer convolutional neural net (CNN), to predict pneumonia using 100,000 chest X-rays using a standard weighted loss function, applying a 99% weight on detecting pneumonia instances and a 1% penalty on incorrectly labeling non-instances as instances.

Given that CheXNeXt was trained based on our diagnostic preferences, one might assume that no other set of preferences would enable the AI to better master our objectives. This supposition is incorrect. In a separate trial, Caplin et al. (2022) train CheXNeXt using the same 100 000 X-rays, here applying a 50% weight on detecting pneumonia instances and a 50% penalty on incorrectly labeling non-instances as instances. Paradoxically, we find that the model that was trained without embedding our preferences does a superior job satisfying our preferences than the model that was trained using our preferences!

The question of precisely why this happens is open and is the subject of ongoing research. One key observation relates to the distinction between the algorithmic loss function used by data scientists to train the model and *economic* incentive to learn. The economic incentive is based on *information content*, or beliefs, rather than algorithmic scores per se. Technically, Caplin, Martin, and Marx (2022) show that the use of highly unequal class weights drives down and systematically distorts the economic incentive to learn. This is an example of the law of unintended consequences!

So what does the AI do differently when its *economic* incentive to learn is so distorted? Our current conjecture is that it becomes *less efficient* at searching through models given the dampened incentives. It is as if it improvements in learning are no longer sufficiently reinforced to effectively guide the direction of search.

We see this result as suggesting a friendly amendment to the current discussion about the *AI alignment problem*. Bostrom's example of an AI that maximizes paper-clip production by turning its programmers into raw materials is an oft-sided if extreme case for putting guard-rails in place to ensure that AI's do not narrowly follow incompletely specified orders. What the cognitive economic approach suggests is that for the AI to satisfy human preferences, one must recognize that the AI has its own cost/benefit calculus that affects what it learns. Accordingly, aligning the AI to optimize human objectives is not accomplished by providing the system with human objectives, but rather by providing it with the correct machine incentives to achieve human objectives.

The obvious challenge our finding raises is how to identify an optimal loss function bearing in mind the impact on AI performance. Intriguingly, the answer can be found by building on existing approaches to the standard alignment problem that *teach* the AI enough about our preferences to not make dangerous mistakes, Ng and Russell (2000) and Hadfield-Mennel et al. (2016) essentially import revealed preference methods of Samuelson (1938) into AI modeling: I cover these pivotal methods in some detail in Chapter 8. Just as the choice environment for a human needs to be manipulated to teach the AI human preferences, so the loss function of the AI needs to be manipulated to teach the human the AI constraints that interfere with learning. The trick is that the optimal loss function to provide the machine may have little to do with the utility function of the human.

5.5 Human-AI Decision-Making Architectures

The third and most familiar place in which humans enter the decision-making pipeline is the point at which a trained model and expert diagnosticians classify cases, and their judgments are somehow combined to arrive at the final classification. In medical diagnoses, a human expert typically gets the final word after receiving multiple machine-generated predictions. Given that they are essentially teamed up with humans in final decisions, the effectiveness of AI-based scores and recommendations is highly dependent on how well humans are able to interpret and utilize them. While Tejeda et al. (2022) find that accuracy improves for many subjects when they receive machine advice, at the same time *algorithmic aversion* makes many reluctant to take improving advice from machines. As noted above, so bad is this in certain medical settings that there are even proposals to not let them both work on the same cases (Agarwal et al. 2023).

Data scientists and subject matter experts are deeply interested in the design of human-machine interactions. In fact there is a whole field dedicated to it called, appropriately enough, *human-machine interaction*. The importance of human judgment has invigorated research on the interaction of human biases with AI performance in cognitive psychology Trueblood, Li and DeLosh (2021). Here what cognitive economics has to offer are conceptual guides as to methods of scoring complete decision-making pipelines. Currently, it is standard to think of the final stage in isolation and to decide between human-in-the-loop and machine-in-the loop architectures. A key part of understanding the best architecture is to conduct experiments to explore how humans can work better with AIs, the subject to which we now turn.

References

Agarwal, N., Moehring, A., Rajpurkar, P., & Salz, T. (2023). *Combining human expertise with artificial intelligence: Experimental evidence from radiology* (No. w31422). National Bureau of Economic Research.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In *Handbook of behavioral economics: Applications and foundations* (pp. 69–186).

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 265–272.

Caplin, A, Martin, D., & Marx, P. (2022). Modeling machine learning. National Bureau of Economic Research (Working Paper 30600).

Christian, B. (2021). *The alignment problem: How can machines learn human values?* Atlantic Books.

Danz, D., Vesterlund, L., & Wilson, A. J. (2022). Belief elicitation and behavioral incentive compatibility. *American Economic Review, 112*(9), 2851–2883.

Hadfield-Menell, D., Russell, S. J., Abbeel, P., & Dragan, A. (2016). Cooperative inverse reinforcement learning. *Advances in Neural Information Processing Systems*, 29.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292.

Ng, A.Y., & Russell, S. (2000, June). Algorithms for inverse reinforcement learning. *LCML, 1*(2), 2.

Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., & Aroyo, L. M. (2021, May). "Everyone wants to do the model work, not the data work": Data cascades in high-stakes AI. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1–15).

Samuelson, P. A. (1938). A note on the pure theory of consumer's choice. *Economica, 5*, 61–71.

Tejeda, H., Kumar, A., Smyth, P., & Steyvers, M. (2022). AI-assisted decision-making: A cognitive modeling approach to infer latent reliance strategies. Computational Brain and Behavior (pp.1–18). *Carnegie-Rochester conference series on public policy*, 33(September), 181–214.

Trueblood, J. S., Li, Y., & DeLosh, E. L. (2021). Human-machine integration in medical decision making. *Current Directions in Psychological Science, 30*(1), 37–43.

Weidmann, B., & Deming, D. J. (2021). Team players: How social skills improve team performance. *Econometrica, 89*(6), 2637–2657.

CHAPTER 6

Work Skills for the Cognitive Economy

In the workplace of the future, most employees will collaborate with AI on a regular basis. It is, therefore, important for economists to understand who benefits from working with AI and why. To remain relevant what is required of humans will be cognitive rather than physical effort. We will need to be able to extract information from multiple sources in real time and make appropriate decisions as a result. On the threat side, AI will supplant expertise that many have spent their entire careers developing, thereby overturning careers and leaving many middle-class professionals with no clear path forward. On the opportunity side, those who have the skills to adapt well to change may find new career paths by boosting the skills that matter as they flexibly incorporate AI into their lives.

Cognitive labor economics addresses these most pressing questions about the future of the workforce in the age of AI. It helps us understand as much as we can about which skills will be devalued, and which enhanced in value. What will the implications be for inequality if we take no proactive policy steps? As discussed in the next chapter it will also be key to developing training protocols for the skills that are required to work effectively with AI in diverse work environments.

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_6

6.1 AI and the Skill Premium

A key economic question about the introduction of machine learning is whether it is a *substitute* or a *complement* for expertise in key tasks. Is algorithmic advice particularly of value for those with high task skills by allowing them to get yet further ahead, or does it particularly helps those with low such skills by lifting base performance? Surely the answer is not one-size-fits-all, but the evidence to date suggests that it may particularly help those who have an intermediate level of expertise close the gap with those at higher levels. The evidence in favor is well-summarized by Autor (2024) in outlining a scenario in which AI might shrink the skills premium and create more solidly middle-class jobs, at least for those with sufficient foundation in basic skills.

The field study of Brynjolfsson et al. (2023) is particularly suggestive. They evaluate the use of generative AI tools that suggest responses to customer service agents. They estimate a significant improvement in productivity overall, particularly among novice workers. These tools allow novices to attain the capabilities of experienced agents in a far shorter time. Quit rates among new agents also fell substantially, due to fewer customer complaints. The AI advice allowed less skilled workers to produce work closer in quality to that of more experienced peers.

A second illustrative study is that of Noy and Zhang (2023), who conducted an online experiment focused on writing tasks. Half of their professional subjects are encouraged to use ChatGPT for writing tasks. The other half are not allowed to use ChatGPT, but rather given access to conventional search engines. Noy and Zhang found significant improvements in the speed and quality of writing output among those assigned to the ChatGPT group. The biggest quality improvements are concentrated at the bottom. ChatGPT closes the productivity gap between good and excellent writers.

6.2 The Importance of Calibrated Beliefs

Tejeda et al. (2022) find that the ability to work well with an AI is far from universal and that there are many cases of algorithmic aversion. A recent study by Agarwal et al. (2023) applies experimental methods to dig into possible sources of poor human-AI collaboration. They conduct an online study with expert radiologists who are asked to identify how likely medical images are to exhibit various well-known pathologies. In some cases their diagnosis is unassisted by an AI, in others they are provided also with a probability assessment generated by an AI assistant. The radiologists are explicitly informed that the AI predictions are well-calibrated, in the sense that when the AI indicates 90% confidence, it is correct about 90% of the time.

The key finding of Agarwal et al. is that AI by and large did not improve the quality of radiologists' diagnoses—even though AI's predictions were more accurate than two-thirds of the specialists studied. So badly did the radiologists perform with the AI in this case that Agarwal et al. suggest that the majority of cases be given either to a human diagnostician without AI assistance, or (more frequently) to an AI without human intervention. There are many reasons to doubt the appeal of such a separation, at least until AI-based diagnoses have become more accepted. There are also issues of unfamiliarity: as AIs get incorporated into the workflow, humans will understand them better. The fact that a generation of radiologists who were provided effectively no training either in probabilistic reasoning or in working with an AI is unable to benefit will not be a durable barrier to integration. In that sense, a key question their research raises is what if anything can be done to train radiologists better.

A key point that Agarwal et al. make concerns the apparent lack of probabilistic sophistication among the diagnosticians. Their failure to integrate AI advice calls into question simple cognitive models of Bayesian updating. In all such models, a diagnostician with relatively low confidence in a particular case would be better off adjusting their assessment strongly in the AI's direction since they were informed that the AI was well-calibrated. That is not what happened in practice. When the AI offered confident predictions, doctors frequently overrode those predictions with their own. There was also a converse issue, albeit less frequent, of diagnosticians replacing their own better predictions with those of the AI even when the AI did not express confidence.

This raises the broader question of how important it is to be aware of one's own skills in order to work well with an AI. Is it possible that being well-calibrated to reality rather than being systematically overconfident or underconfident is a general purpose skill that impacts a worker's ability to benefit from AI advice? Those who are overconfident in their own abilities may rely exclusively on themselves even when incorporating what the machine indicates would be beneficial. This appears to have been the case in the experiments of Agarwal et al. Conversely, those who are underconfident may defer too much to the AI even though they have much of value to contribute. Pushing one stage further, they may even stop learning because they defer too much to the machine.

There is already much interest in calibration in the machine learning community. When algorithmic advice is passed on to human decision-makers, it is seen as important to convey it in the language of *probabilities*. That is why Agarwal et al. make clear that machine scores are well-calibrated and hence can be interpreted probabilistically. The importance of translating algorithmic scores into probabilities is what has driven interest in calibration within the machine learning community. So important is this felt to be that when algorithmic scores are poorly calibrated, it is standard practice to *recalibrate* them into probabilistic form. This is what underlies the industry standard methods that we used in the wisdom of the crowd context to recalibrate the subjective beliefs of members of human labeling communities in the last chapter. We borrowed a page from their book in that case.

6.3 The ABCs of Who Benefits from Working with AI

I now outline a first experimental investigation of the impact of calibration on the ability to work effectively with algorithmic advice in Caplin et al. 2024. The key innovation is that we use the experimental lab to estimate individual calibration without AI as well as skill. We then investigate how this estimate of calibration skill impacts value-added with the AI. Our design also allows us to explore the important links between ability at a task and AI value-added.

Our starting point is the emerging consensus that AI helps those of lower ability close the productivity and earnings gap with those with higher levels of ability and the suggestive evidence that those whose beliefs are poorly calibrated make poor use of AI advice, both discussed above. In our research, we design an experiment to answer two obvious follow-up questions. Are there important *individual differences* in how well beliefs are calibrated that impact value added of AI advice? Given that ability and calibration are likely to be correlated, how does this relate to compressing the ability-based wage gap?

Our experimental design is close in spirit to that of Agarwal et al. (2023). It shares the feature that subjects make incentivized probabilistic assessments in classifying images. But their medical classification task is both time consuming for participants and costly in terms of subjects, who are highly paid medical specialists. The limited sample size precludes them from digging into individual differences. We use a far simpler task that is suitable for implementation in Prolific since it does not require specialized training. Our task is to classify how likely it is that the subject in a photograph of a face was under 21 years old at the time the image was taken. We are able to gather 160 classifications for all subjects in our Prolific sample of some 1500 subjects. Given that this task has not previously been studied experimentally, we run a control group that never get AI advice to monitor and account for possible dynamic effects due to increased task familiarity and or tiredness. When estimating the treatment effect, it is important for us to make allowance for measurement error, since we find task skills and calibration are correlated. We adapt the *Obviously Related Instrumental Variable* approach of Gillen et al. (2019) to our setting to estimate treatment effects.

The results are now in and provide clear answers to all questions. We find stable individual differences in both how skilled and how calibrated subjects are. There is a positive correlation: those who are more highly skilled are generally better calibrated. We also find that the group that benefits most from working with the AI are those of lower ability whose beliefs are well-calibrated. Those who are well-calibrated have higher value added with an AI either than the overconfident majority or the under-confident minority. The obvious follow-up questions are how generalizable calibration skills are across tasks, and how teachable these skills are. This is a key subject of ongoing research: stay tuned for updates.

6.4 Algorithmic Awareness

There is an obvious converse of being well-calibrated, which is how well human decision-makers are able to understand the strengths and weaknesses of the algorithms they work with. We refer to this as *algorithmic awareness*. The broad idea that the human model of the AI matters is strongly suggested at the informal level in cognitive psychology. The challenge is how best to operationalize and quantify skills related to interpreting and complementing algorithmic advice. I am aware of no research that addresses these forms of skill, which I believe may be among the most important of all for the algorithmic age. To understand why let me list a few qualitatively different responses humans may have when working with an algorithmic adviser.

1. *Defensive Dismissal*: One way to respond to an algorithm is to dismiss it as "crazy". Those who are so inclined will doubtless be able to point to cases in which algorithmic advice is absolutely ridiculous and reflects inference that essentially any human could best. Algorithms for image classification work at the pixel level. This means that the algorithm is always ready to determine how likely a picture of Mickey Mouse is to be cancerous as opposed to a non-cancerous blood cell. The folk theory of algorithmic aversion is precisely that seeing obvious mistakes produces mistrust that is globalized. Of course, the theory has not been rigorously tested. In defense of our algorithmic advisers, if they could get together and talk, what do you think they would be saying about us? Not very good adding machines, are we?
2. *Defeated Acceptance*: It is likely in future that we will find out that algorithms are, on average, superior in performance to most of us even if we put in full effort. There may be those who simply give up or at least reduce effort based on algorithmic advice lowering value added.
3. *Best of Both Worlds Independence*: Rather than ignore or blindly accept algorithmic advice, it would be better for human experts to apply their own unique perspectives to get a read on the likely diagnosis in addition to taking in the algorithmic advice. They might then be able to taking advantage of the cases in which the algorithm is better, while always incorporating any unique information they can glean.
4. *Algorithmic Sophistication*: If an individual really gets an understanding of the strengths and weaknesses of their algorithmic advisers, they can adjust their learning to focus particularly on areas in which the algorithms are weak. In this manner, they may actually train differently and work to develop a form of comparative advantage in learning that makes them particularly complementary to the algorithms they work with on a daily basis. This is a very deep skill if it can be honed and is on a high order of understanding oneself, understanding the machine, and knowing which aspects of what the machine does not understand are possible to learn. Would

top chess players who have been trained using AIs know enough about AI weaknesses to know when to overrule their advice and thereby beat the AI?

Operationalizing the above concepts is a first order cognitive economic challenge. Ideas are welcome.

6.5 Adaptability, Resilience, and Search Skills

Like it or not, the coming of the cognitive economy will lead to radical shifts in the types of jobs that are available and in the skills required to carry them out effectively. To survive in an era of job market turmoil requires various forms of adaptability and resilience that have been little studied. Interestingly this links with the one area of labor economics in which it is understood that cognitive constraints matter: the theory of job search. Decisions on whether and how to search for jobs, including when to quit, what to do in the face of an impending layoff, when to take time out of the labor force and retool, and when to retire all take cognitive effort. The essentially cognitive nature of these questions is implicit in the basic economic model of job transitions, which is based on the theory of search.

While early conceptions of search costs were physical, e.g. going to an interview, in later conceptions it became clear that the constraints were more cognitive than physical in nature. No one can possibly take in all job options all of the time. What matters is how close people come to having a full understanding of what their outside options might be at all times, in order to make appropriate choices of when to negotiate with the current employer, when to search intensively, and when to move to a new employer.

As always in cognitive economics, there are major challenges of measurement. Administrative data on earnings show that job switches play a key role in earnings, with some moving up the job ladder, and others staying in place or even slipping back. Yet administrative data alone have little to say about why some move up the job ladder and why others do not, and the extent of this depends on differences in standard work skills as opposed to skills in identifying and pursuing opportunities. Is it largely that better workers have better outcomes in all job phases than worse workers, or is a significant portion determined by the differential abilities of similarly skilled workers in recognizing and taking advantage of potential opportunities? Are there significant individual differences in how well tied to reality are the beliefs that different individuals hold? Is there a simple cognitive instrument that might identify such differences? There is no reason to expect the required forms of adaptability and resilience to be uni-dimensional. Elements may include threat identification (e.g. impending layoff); opportunity identification (e.g. searching while on the job); realism (e.g. accurately anticipating contingencies); flexibility (e.g. adjusting job search strategy in light of experience); and adaptability (e.g. taking time out of the labor force to train in newly important skills).

6.6 Beliefs about Future Earnings

In addition to using cognitive instruments to measure the skills that are required to make successful job transitions, we need to understand why some people see it as worthwhile to actively search for jobs while others do not. To a large extent this depends on their beliefs about contingent future earnings: will their wage go up if they stay in their current job, or do they think they will need to move to advance up the job ladder? Recent advances in understanding in this regard relate to improved measurement of the subjective beliefs about future earnings that guide labor market decisions.

Survey measures of probabilistic beliefs about future earnings were pioneered by Dominitz and Manski (1996). They elicited probabilistic estimates of the cumulative distribution of total household income, before taxes, over the next 12 months. After posing questions on the maximum and the minimum possible values, they asked for several points on each respondent's subjective cumulative distribution. There have been many advances in measurement since that time. In an ingenious effort to capture folk ideas in a quantitive manner, Delavande and Rohwedder (2008) introduced a design in which the range between stated minimum and maximum is visually divided into a number of equally sized bins, with respondents placing balls into the bins to reflect their belief in how relatively likely are the corresponding ranges. What makes this design compelling is that visual devices are perhaps more intuitive for many than are strictly quantitative questions.

A recently initiated panel dataset, the Copenhagen Life Panel (CLP), provides state-of-the-art measures of subjective expectations of future earnings from work designed precisely around the question of why some quit and others do not. It is a contingent survey of precisely the form that cognitive economics demands. It focuses on beliefs about the impact of job transitions, both quits and layoffs (Caplin et al. 2023). For those working for pay, measures are elicited of the probability of continuing in this work all year, quitting during the year, and being laid off during the year. In all conditions, probabilistic questions in the balls-in-bins style are asked about future incomes. The sample is drawn from and linked to the Danish population registries. This allows us to conduct many credibility checks on these answers, which are passed with flying colors. Having such a linked panel in which we observe both beliefs and outcomes is key to finding out what is and is not understood by workers.

Research using the CLP to understand job transitions is of the highest promise, yet at the very earliest of stages. Going forward, the ability to track both beliefs and outcomes provides a solid groundwork for understanding the transition to the cognitive economy. Research along these lines is outlined in Chapter 9.

References

Agarwal, N., Moehring, A., Rajpurkar, P., & Salz, T. (2023). *Combining human expertise with artificial intelligence: Experimental evidence from radiology* (National Bureau of Economic Research Working Paper 31422).

Autor, D. (2024). *Applying AI to rebuild middle class jobs* (No. w32140). National Bureau of Economic Research.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work* (National Bureau of Economic Research Working Paper 31161).

Caplin, A. (2023). *The science of mistakes: Lecture notes on economic data engineering*. World Scientific Press.

Caplin, A., Deming, D.J., Li, S., Martin, D.J., Marx, P., Weidmann, B. and Ye, K.J., 2024. *The ABC's of Who Benefits from Working with AI: Ability, Beliefs, and Calibration* (No. w33021). National Bureau of Economic Research.

Dominitz, J., & Manski, C. F. (1996). Eliciting student expectations of the returns to schooling. *Journal of Human Resources, 31*, 1–26.

Delavande, A., & Rohwedder, S. (2008). Eliciting subjective probabilities in internet surveys. *Public Opinion Quarterly, 72*(5), 866–891.

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192.

Tejeda, H., Kumar, A., Smyth, P. and Steyvers, M. 2022. AI-assisted decision-making: A cognitive modeling approach to infer latent reliance strategies.

Computational Brain and Behavior, pp.1-18. *Carnegie-Rochester conference series on public policy*, 33(September), 181–214.

CHAPTER 7

Cognitive Economics of Teaching

Cognitive economics is the science of mistakes. As such it applies to teaching that is designed to reduce mistakes. Following Bloom's taxonomy, we want students not only to learn facts, but know to know how to apply them to questions of interest, and how to identify patterns and relations among different branches of knowledge. It is these higher level decision-making skills that may be most important in defining the forms of human capital that will be of highest value as we transition to a cognitive economy.

In this section I make a few observations on teaching higher order skills of importance in decision-making, as well as on how to test for the various levels of understanding. By way of illustration I consider teaching of *calibration* that builds on its importance in working with AI discussed in the last chapter. One virtue of this is that there is already profound interest and some relevant research. I outline psychological research that is suggestive about the possibilities for teaching calibration. But this research is neither definitive nor quantitatively compelling. We urgently need to better identify and teach calibration and other skills for the cognitive economy. We also need to better evaluate how well these have been internalized by students, and I follow up in this regard on an important and under-studied proposal on how to design tests to be more informative about the skills that matter.

I also take up the issue of what students know about the future rewards of the courses of study they select. The evidence suggests that they know

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_7

very little, and much of what they think they know is inaccurate. Research into how best to communicate this information effectively is at the earliest stage.

I close by proposing courses in cognitive economics at all levels, starting with high school. In the long run, it is profoundly important to have a social scientifically literate population. At present, there is no country in the world in which this is true. Worse still, the supply of teachers who have requisite skills is desperately limited. Changing this should be a super high priority for social progress. I keep discussion short at this stage because the teaching industry appears to be a closed shop. I would be more than delighted to have the opportunity to revisit these matters at a later date.

7.1 Teaching Calibration

As indicated in the last chapter, initial research suggests that it will be of great value for workers of the future to understand their own skills well enough to make good use of algorithmic advice. As further noted, a particular type of such skill is the ability to assess uncertainty in a manner that is linked with external reality rather than being naively optimistic or defensively pessimistic about one's ability in this regard. The obvious question is how to design methods of teaching to improve the kinds of realistic assessment of abilities of self and others that will be required in the workplace of the future.

Clear examples of why this matters are to be found in the medical profession. Studies reveal a profound lack of probabilistic sophistication. Many medical professionals are unfamiliar with even the basic rules of probability. If anything, the medical profession rewards a sense of certainty. This is terrible. Accurately assessing beliefs is something that diagnosticians themselves should be trained in from an early age. The need to build up such sophistication will likely grow when future pathology occurs within the context of AI systems, as is sure to happen. At this point it will be important to conduct experiments with medical trainees to explore how AI advice impacts (1) perceptual learning and (2) the development of probabilistic competence.

Thinking more broadly, there are many open questions related to calibration as a trait. To what extent is it domain general? To what extent does an overoptimistic assessment of abilities operate across different forms of interaction? To the extent that it does not operate universally, is it nevertheless fixed in important work arenas, such as medical diagnoses for subject matter experts? The second and more fundamental open question is whether calibration is a teachable skill. My own belief is that it must indeed be with sufficient feedback and testing.

Psychologists have asked these questions and provided isolated answers, but this research has not elevated to a general level. Lichtenstein and Fischhoff (1980) suggest that calibration skills are to some extent general and to some extent teachable, but no general teaching tools have been developed. In their informal small scale investigation, most of the improvement happened quickly, with a total intervention time of some 60 minutes. People improve in part because they decrease their use of 0% or 100%. Their findings also suggest that the improvements in skill are somewhat generalizable across tasks, i.e. if you get trained to be better calibrated on task X, there's often an uptick in accuracy on task Y. The larger question is wide open, particularly after some less than fully encouraging findings of Moore et al. (2017).

Broader tests are currently being implemented in ongoing cognitive economic research. There are also open questions about what forms of teaching methods are worth exploring. Intriguingly, one possibility is that simply working with an AI decreases the use of extreme answers. In ongoing research we have early indications that such an effect might be present. An obvious rationalization is that the AIs never display 100% confidence, despite on average being better than most humans. This may push people off their claim of certainty. In addition to implementing teaching modules experimentally the ideal in research on how to train calibration will be to develop corresponding modules for implementation in the Danish population registries, where effects can be followed over the course of the subsequent years.

7.2 Testing Properly

When we design methods to teach, we need also to develop complementary grading protocols. A grading scheme shapes both the incentive for students to learn and what can be understood by teachers about how well they are conveying information and how well students are taking it in. Current practice in multiple choice tests is extraordinarily primitive in this regard. A question is posed. Five possible answers are listed. One and only one is correct, or at least deemed to be correct. Grades depend on the number of such correct answers. Once each question has been graded the number of correct answers is added to produce an overall score. At this point the test is discarded and the student either moves on or is turned back and asked to repeat the class and the test as a function of performance. A moment's thought reveals how much this method sweeps under the rug and how much more could in principle be revealed even within the constraint that one must stick with a multiple choice format.

A question might occur to the engaged reader. If forced choice protocols are beaten by elicited belief protocols in wisdom of the crowd labeling, might they not also be worth trying in testing and training protocols? If you did think this, welcome to the club. But I would like you to know that neither you nor I were first to pose this question. For those who like pedigree, the high value of eliciting subjective beliefs about answers to test questions was first highlighted by Savage (1971) and De Finetti (1965). Together they are largely responsible for placing subjective probabilities at the center of economic and social thought, a revolution in thought whose importance can hardly be overstated. It is absolutely fundamental to modern social science. Savage also pioneered analysis of the proper scoring rules that incentivize truthful revelation of subjective beliefs that played such a key role in the last chapter. To square the circle, he studied them precisely *because* he wanted to see them introduced into the grading process. Since this pioneering work a small but important psychological literature has developed on probabilistic scoring of multiple choice tests, but has had little impact on the educational establishment.

While not broadly adopted, proper scoring rules have been applied in at least one field, *decision analysis*. Bickel (2010) indicates a number of important respects in which tests graded with proper scoring rules reveal richer information about what is understood. For example, there are answers that many would have got correct if forced to pick one response, but about which there was high uncertainty, calling for the topic to be revisited. He also finds that calibration varies across subjects. Averaging across students in his courses, he finds that answers were approximately calibrated: the answer 0.9 corresponded to roughly 90% of the corresponding answers being correct. Yet, he noted large individual differences in how well-calibrated students were. There were students who did not approach the maximum degree of certainty, yet earned some of the highest marks because they assessed their more limited state of knowledge well. Conversely, there were students who expressed great confidence in answers being correct and were therefore penalized heavily because too high a proportion of such answers were wrong.

It should go without saying that much broader consideration of proper scoring rules and other changes in test design are warranted. Likely, there are those who will claim that proper scoring rules are "too complex". One can even imagine jokes about the course that students would need to understand them. My own experience teaching suggests exactly the opposite. If there are clear rules about how to earn a grade, students follow them in fine detail. Grades are so profoundly important that how they are earned, which varies course by course, is job number 1 to explain. If proper scoring rules of a few well-tested forms were to be introduced and used in multiple settings, students would understand them intimately. Online resources would explain them engagingly. All that is required is a will and an appropriate research effort. Seventy years is a long time to wait to even consider a hugely important idea about teaching. Time to get it scientifically evaluated.

There is a link between proper scoring rules and calibration that was first noted by Savage (1971) when he proposed their use. These scoring rules may be of particular value when used in training calibration. Surely one would expect students who were punished in exams for their overconfidence to learn to temper their judgment in the direction of realism. *Reinforcement learning* alone should accomplish this. In this manner, teaching of calibration may be strongly linked with the use of proper scoring rules in testing protocols introduced directly above.

If this thought occurred to you, I would like to note again that neither you nor I were first in this regard. It is precisely this feature that was viewed by Leonard Savage as the key advantage of using proper scoring rules in testing. He felt that failure to reward calibration and moderation of beliefs are critical problems that the use of proper scoring rules might help solve. It is past time to put Savage's vision to the test.

7.3 Teaching Students the Value of an Education

An important early application of survey-based probabilistic measures of future earnings relates to the value of an education. Dominitz and Manski (1996) provided early evidence that student expectations of returns to schooling impact their educational choices. Recent years have seen this issue investigated in greater depth. Arcidiacono et al., 2014, pose survey questions on beliefs, stated preferences, and probabilities of choosing particular occupations of undergraduate students at a well-known university. They find large differences in expected earnings across occupations, and substantial heterogeneity across individuals in the corresponding ex ante returns. They find that many individuals are willing to give up substantial amounts of earnings by not choosing their highest-paying occupation. Wiswall and Zafar (2021) study how individuals believe human capital investments will affect their future career and family life. They find evidence of students sorting into majors based on perceived ex ante returns. Family expectations are found to be particularly important for females' major choices. In a follow-up survey conducted six years after the initial data collection, they find a close connection between the expectations and realizations.

Given that expectations clearly play into educational and career decisions, it is important to know how well-grounded in reality these expectations are, and where poorly founded to try to correct them. Conlon (2021) finds evidence of significant misinformation about the earnings associated with different majors, and shows that providing accurate information significantly changes major choice in the corresponding direction. This points to an important cognitive economic project to measure students' perceptions of the skills required for the cognitive economy, as further discussed in Chapter 9.

7.4 Teaching Cognitive Economics

The lack of scientific interest in social and economic questions I see around me has been endlessly frustrating. Many hold strong opinions about the right thing to do on social issues of fearsome scientific complexity. False certainties divide us into social clubs whose dividing walls are made all the stronger by the weakness of their foundations. I am a member of the *Non-Conformist Club* as founded by Tony Hancock (with a tip of the hat to Alan Galton and Ray Simpson). We need more adherents. This book is my best shot at recruitment.

The best way to get others to retreat from false certainties to scientific curiosity would be to introduce social scientific training from the youngest of ages. Courses in cognitive economics would fit the bill. Just once in their lives, students should be taught that one can reason about individual and social outcomes rather than take sides in debates built on foundations of quicksand. Such a course would not only provide a common language for scientific discussion of human affairs, but also would provide all of us with the profound and novel insights that younger thinkers and, perhaps more importantly, older thinkers new to this class of question can have.

Of course, we also need course materials for all levels of cognitive economics, a new major, etc. That will all have to be coordinated with scientific advances and the gradual transformation of the social scientific landscape to a more open architecture. Development of such materials will be a key charge of the Accelerator for Cognitive Economics introduced in Chapter 10.

References

Bickel, J. E. (2010). Scoring rules and decision analysis education. *Decision Analysis, 7*(4), 346–357.

Conlon, J. J. (2021). Major malfunction: A field experiment correcting undergraduates' beliefs about salaries. *Journal of Human Resources, 56*(3), 922–939.

De Finetti, B. (1965). Methods for discriminating levels of partial knowledge concerning a test item. *British Journal of Mathematical and Statistical Psychology, 18*(1), 87–123.

Dominitz, J., & Manski, C. F. (1996). Eliciting student expectations of the returns to schooling. *Journal of Human Resources, 31*, 1–26.

Lichtenstein, S., & Fischhoff, B. (1980). Training for calibration. *Organizational behavior and human performance, 26*(2), 149–171.

Moore, D. A., Swift, S. A., Minster, A., Mellers, B., Ungar, L., Tetlock, P., Yang, H. H., & Tenney, E. R. (2017). Confidence calibration in a multiyear geopolitical forecasting competition. *Management Science, 63*(11), 3552–3565.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association, 66*(336), 783–801.

Wiswall, M., & Zafar, B. (2021). Human capital investments and expectations about career and family. *Journal of Political Economy, 129*(5), 1361–1424.

CHAPTER 8

Cognitive Economics Takes Off

The origins of cognitive economics date back many decades. In this chapter, I provide an overview of these origins, why development has been so long delayed, and what has recently happened to speed progress. I organize this overview around Rostow's powerful schema defining five distinct *stages of growth* for economies in the process of economic development (Rostow, 1959). I see these stages as being at least as applicable to the growth of a new scientific discipline as they are to the growth of an economy. Furthermore, I draw a parallel with the stages of growth in data science, which offers a more contemporary analogy.

Rostow's stage 1 is traditional society: for cognitive economics, that is traditional economics and cognitive psychology living in their separate academic silos. This stage is akin to the early days of data science, where disparate fields such as statistics, computer science, and domain-specific knowledge existed separately, each advancing independently without much cross-pollination.

Rostow's stage 2 is the period in which preconditions for takeoff to sustained growth are put in place. For cognitive economics, this involves the initial interdisciplinary collaborations and the development of basic models and theories that bridge the gap between economics and cognitive psychology, In the context of data science, stage 2 represents early integration of statistics, computer science, and domain-specific knowledge, leading to the creation of foundational tools and techniques.

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_8

Stage 3 is the takeoff itself. In cognitive economics, this stage is well under way with a surge in research activity and application as foundational theories prove their worth and attract wider interest. While on an altogether smaller scale at this point, this mirrors to some extent early growth in data science driven by advances in computational power, ever growing availability of massive datasets, and the development of sophisticated algorithms with demonstrated power to achieve key goals, such as accurately reading messy hand-written addresses.

Of course, it is never quite as simple as 1, 2, 3 either for an economy or for a science. In most cases, there is what might be called a stage 1.5, defined by discontent with the status quo and the build-up of pressure for change. Discontent is a powerful motivation for doing the hard work of building preconditions for takeoff. A precondition for preconditions, as it were. This stage has been particularly important for cognitive economics, as detailed below. In data science, this phase was marked by growing frustration with the limitations of traditional statistical methods and the inefficiencies of handling large datasets, which spurred the demand for more integrated and powerful analytical approaches.

Rostow's stage 4 is the drive to maturity, and stage 5 the age of mass consumption. Modulo a switch in the interpretation of stage 5 to mass application, I think this provides an apt framework for conceptualizing the future growth of cognitive economics. The drive to maturity in data science corresponds to the refinement of methods, the establishment of best practices, and the standardization of tools, leading to the widespread adoption of data-driven decision-making across industries. For cognitive economics, this stage involves the consolidation of research findings, the development of comprehensive models, and the integration of cognitive insights into mainstream economic analysis. Stage 5, the age of mass application, parallels the pervasive influence of data science in every facet of modern life, from business to healthcare to governance. Similarly, cognitive economics will reach its zenith when its principles are ubiquitously applied to shape policies, business strategies, and societal norms.

The parallels with data science are not random. In fact there are reasons to see the developments as complementary. It is the explosion in the use of AI that adds urgency to progress in cognitive economics. Were this some other new branch of social science, perhaps our stately pace would be adequate. But what I am discussing instead is a form of social science that will be of most value in the transition to a cognitive economy. We will need to ramp up our speed by several orders of magnitude if we are to keep up as AI sweeps the economy and potentially eradicates many of the jobs that have traditionally supported middle-class lives. This will happen at the speed of data science rather than social science. There really is no time to waste.

In the rest of this chapter, I expand on the stages of growth of cognitive economics, up through where we stand today, which is at the beginning of stage 4, the drive to maturity. This is a natural lead into the final chapter of the book in which I make proposals for accelerating through this stage to bring about stage 5, the age of mass application.

8.1 Stage 1: Samuelson and Revealed Preference

If I had to pick a date when the need for cognitive economics was first realized, it would be 1959. To provide context, I want to highlight the profound insight of Paul Samuelson in 1938 when he introduced *revealed preference* theory into social scientific thinking. I see this as a methodological and substantive milestone, spiritually aligned with transformative results in other sciences, such as Bell's theorem in physics. At its core, Samuelson's method catalyzed the development of cognitive economics. Let me outline his foundational contribution and its enduring significance.

As a spectacularly able student of economics, Samuelson was well-trained in classical utility theory. This theory had been introduced in the middle of the nineteenth century to help explain why diamonds are expensive and water cheap, despite the infinitely greater value of the latter when thirsty. The answer to this *diamond-water paradox* suggested by utility theory rests not on *total* utility but rather on *marginal* utility. Sure, if one had to pick between *either* having access to water *or* having a diamond, the choice would be clear. But that is not really the situation we face in a world in which water is plentiful and diamonds scarce. On the margin, more water was worth relatively little in the Europe of that era, since it falls from the sky, while diamonds were and still are scarce and highly valued on the margin.

Fast forward fifty years to Vilfredo Pareto, who noticed that choice is not impacted by any relabeling of utilities that preserves order. Better is better. But to say "more better" is a bridge too far. One can stretch and contract the utility scale so as to render marginal utility meaningless. So in the (then as now) modern theory of rational choice, utility functions are replaced by more abstract preference orderings. Meaning is given to strictly preferred to, strictly less preferred than, and indifferent to. In preference theory, unlike in early utility theory, no meaning is given to any numerical utility scale that might previously have been thought to register utility differences. Marginal utility was dead.

Slow forward another forty years or so to Samuelson. He was well aware of the critique of marginal utility due to Vilfredo Pareto. But he did not find it scientifically compelling due to its side-lining of measurement. What struck him was how thoroughly *non-operational* the final theory of rational choice seemed. No claim was made that preference relations are observable. Rather the theory was that these unobservable relations determined choice. Given options, the decision-maker is hypothesized to pick an option that is preferred at least weakly to all alternatives. But where's the evidence? A choice is a choice is a choice for sure. But to "explain"choice as resulting from a process of optimization seems circular. To rationalize choice of an apple over an orange, one must model the former as being preferred. When asked to go deeper, it might be tempting to say that preference is revealed by the observed choice.

Samuelson's complaint was direct and blunt.

> The discrediting of utility as a psychological concept robbed it of its only possible virtue as an explanation of human behaviour in other than a circular sense, revealing its emptiness as even a construction. (Samuelson, 1938, p. 61)

He then posed the million dollar question: What can stop rational choice theory from being tautological? Is there any pattern of behavior that it *cannot* rationalize?

What a wonderful question that turned out to be. To address it, Samuelson invented a method of thought. This method required him to think creatively about the *datasets* that we use to operationalize economic constructs. Measurement protocols belong front and center in *defining* model constructs. To break the trap and render rational choice theory testable, he insisted that we would need far richer data on behavior than is traditionally available. What we would need to do is to vary the choice context and look for structure in the resulting behavioral data that would either show us patterns in preferences that were consistent with classical rational choice theory or perhaps tell us that this is critically incomplete.

The ideal form of data that he specified is now seen as the most standard form of economic data: choice among affordable options from all conceivable budget sets.

> All that follows shall relate to an idealised individual *not necessarily, however, the rational homo-economicus*. I assume in the beginning as known, i.e. empirically determinable under ideal conditions, the amounts of *n* economic goods which will be purchased per unit time by an individual faced with the prices of these goods and with a given total expenditure. It is assumed that prices are taken as given parameters not subject to influence by the individual. (Samuelson, 1938, p. 62: my italics)

Here is the key innovation. Samuelson challenged traditional, non-empirical approaches to utility theory by proposing a method to vary choice contexts experimentally. This approach, insisting on detailed empirical data collection, provided a testable framework for rational choice theory, a significant leap from theoretical postulates to actionable empirical research.

It is hard to explain how profound an idea Samuelson introduced. Like essentially everything that is most creative in social science, it is a method of thought rather than a particular hypothesis. For many economic theorists, revealed preference theory of this kind is something of a religious exercise undertaken as one technical exercise among many. I think instead that it is a dramatically different method of thought. Choice data is the basic given. Maximization of preferences is one possible model of these observations, albeit an important one. There are other theories that might better explain the data. In fact Manzini and Mariotti (2007), introduce a weakening of the strong axiom of revealed preference that characterizes an important class of models with bounded rationality. Subsequent research along these lines is contributing importantly to the growth of cognitive economics. As part of this research path there is also great interest in further expansions of ideal data. It remains in many ways shocking to me that we have taken so long to pick up on Samuelson's implicit call for richer modeling of choice behavior.

From the perspective of cognitive economics, what Samuelson is saying is that we need to think long and hard about what our abstractions are meant to imply for data. Not the data that we are currently gathering, but rather *ideal data* that springs from our scientific imaginations. His contribution was to imagine what could be learned if we were to gather ideal data in a form that had never really been gathered. What his work shows is that choice from all conceivable budget sets as prices and incomes vary is conceptually ideal, leaving applied researchers to try to approximate it as best possible in realistic settings and/or in experimental data. There is now a large body of work in economics based on gathering ever closer approximations to Samuelson's ideal. Some of the key contributions in this literature relate to patterns of spending over the life cycle (Adams et al., 2014; Blundell et al., 2008). Not a bad day's work.

At this point, the analogy with early data science becomes particularly powerful. Samuelson's pioneering efforts not only reshaped economics but also exemplified a shift toward empirical rigor that would later become a hallmark of data science. His emphasis on rich, structured data collection mirrored the transformation in data handling and analysis that data science would eventually champion. Just as Samuelson called for a transformation in the collection and analysis of economic data, early data scientists were pushing for a revolution in handling and analyzing large datasets. They developed new computational tools and statistical methods to manage and derive insights from data that were previously unmanageable. This methodological evolution in both fields represents a parallel shift from abstract theories to data-driven practices that could be rigorously tested and applied.

Further to the point of the key contribution being methodological, it is notable that Samuelson did not get the testable conditions for preference maximization quite right. He got part way when he established that failure of the *weak axiom of revealed preference* would render the standard theory of rational choice false. According to this classical theory, any unchosen yet affordable combination of goods must have been rejected in favor of what was in fact chosen. Hence if some pair of observed choices could be switched between budget sets and total costs lowered, that would contradict classical maximization logic.

So far so good. But Samuelson had higher ambitions. His goal was not only to find conditions on data that are *necessary* for standard rational choice theory to apply but to close the loop by establishing conditions that are *sufficient* for the theory to apply. In other words, a faithful translation of the theory of rational choice to a condition on ideal data. Where he stopped short is that he did not find the full characterization. It was left to others to close the circle. We know now that a strengthening of the weak to the *strong axiom of revealed preference* is not only necessary for the classical theory but also sufficient. The necessity part is intuitive:

the same logic that rules out switching a pair of choices implies that there can be no cycles of strict affordability, since if there were, the same goods in total could have been bought for less. After dotting i's and crossing t's, that is the strong axiom of revealed preference.

A plodding interpretation of what Samuelson accomplished would see him as having imposed on economics the religious principle that choice reveals preference. Rejection of this religion is then taken to be a sign of being less beholden to the rationalist ideology than him. This is a complete misunderstanding. Not for one minute did Samuelson "believe in "utility maximization or intend to start a religion for true believers. His intent was entirely opposite: to enrich data to the point that it can be seen as a theory that might possibly be rejected. So we should go ahead and gather that data, reject the theory where necessary, refine it where otherwise, and advance the science of human behavior.

8.2 Stage 1.5: The Winter of Cognitive Economic Discontent

What Rostow did not formalize in his description of stages of economic growth are the drives that motivate moving away from tradition. In the case of cognitive economics there was an explicit stage of discontent with the status quo, stage 1.5, that motivated working toward change. This discontent had two basic sources. First, the strong axiom of revealed preference is often violated in laboratory and field data, just as Samuelson knew it would be. For example, individual choice among a fixed set of alternatives is often random, which in the basic story implies the absurdity that everything is indifferent to everything else. The second source of discontent is more theoretical than empirical. From the modern viewpoint, possibly the most important limitation of classical choice theory is the underlying assumption that information is perfect. We no longer model most choices as fully informed, and the standard dataset is blatantly insufficient for identification when information is imperfect. With incomplete information it is easy to create the kinds of cyclic choice patterns that the strong axiom of revealed preference rules out.

The fact that information is imperfect and the resulting difficulty in inferring preferences from choice behavior alone is precisely the backdrop to much of the research in cognitive economics laid out above, particularly in Chapter 3 on complex choices. That chapter in fact opens with the observation that when information is imperfect, which is always, the simple equivalence between preference and choice breaks down. Standard choice data reflects both what people like and what they understand. So a scientific observer of choice in complex environments is left with a conundrum in that they cannot establish one way or another how well or poorly informed the decision-maker was. Many of the data innovations in the book are responsive to this precise challenge.

With this by way of background, let me explain what happened in 1959 that might have led to the immediate launch of cognitive economics. It was in that year that Henry Block and Jacob Marshak first laid out the essential measurement challenge in cognitive economics: identifying new forms of data that might reveal cognitive constraints. They were motivated in this by their study of randomness in choice, which they noted contradicted the strong axiom of revealed preference. To allow for this data pattern they modeled the forces that cause people to make different choices from one and the same choice set. They borrowed many of their ideas from cognitive psychologists, who based theories of such randomness on imperfect perception and cognitive constraints. In fact in key respects, the mathematical psychologist Duncan Luce had already scooped them in developing the *logit* model of random choice (Luce, 1958). In turn he was building on a truly venerable tradition in experimental psychology dating back to Weber (1834), and the psychometric curves of Fig. 3.1 in Chapter 3.

There is a key distinction between the psychological and economic approaches to random choice. Building on the psychological tradition, Luce was thinking in terms of imperfect perception and called his stochastic choice function a *discrimination structure*. By way of contrast, when Block and Marschak adapted psychometric ideas to the economic context, they were building on the revealed preference logic of Samuelson. They connected randomness in choice both with randomness in perception and with randomness in preferences, which for historical reasons they referred to as randomness in utility. Their *random utility model* has subsequently been developed into a centerpiece of applied microeconomics, defining the field of discrete choice. But economists appear to have forgotten the model's cognitive underpinnings, and by and large ignore perceptual issues entirely. In most models that allow for random utility, it is simply *assumed* that information about options is perfect. With this, one can estimate how valuable is each of the characteristics that attract consumers to buy one good over alternatives. Confusion related to limited perception of complex options is typically ignored.

Unmeasured constructs can't talk I guess. We are in the early days of rectifying this omission, with Manzini and Mariotti (2007) taking the lead in developing models of stochastic choice based on incomplete examination of available options.

This neglect of cognitive constraints is not at all what Block and Marschak had in mind. As was Luce, they were very much focused on cognitive constraints. Their primary example of stochastic choice is entirely perceptual: a wholesaler's repeated choice between two 10-ton carloads of the same merchandise in which neither quantity nor quality can be perfectly ascertained. They thought of both perception and utility as having random elements. They realized that when two quite different and unmeasured sources of randomness get funneled into one and the same choice, it is essentially impossible to assess their relative importance. Hence their economic twist introducing random utility into what had been a purely perceptual story came at a scientific price, as they themselves pointed out. In fact they recognized the depth of the problem this created. They were insistent that no form of standard choice data would be adequate to separate "information and desirability". As a result they never claimed to identify utility in the classical sense but rather some mixture of utility and perceptibility:

> All of the various definitions of utility given in this paper will be related to the empirical entities, called "alternatives". Each of these is identified precisely, but combines the information and the desirability aspect in some unknown though presumably not too changeable fashion. (Block & Marschak, 1960, p. 175)

Noting the depth of this and other essential identification problems we face in social science, Block and Marschak called for new forms of measurement to separate out latent forces that are hard to identify in classical forms of data. In the case of cognitive economics they called for data that would allow researchers to separate out utility from cognitive limits. They encouraged the search for what they called "new basic observations". Much of the work in cognitive economics is based on just such innovations.

8.3 Stage 2: Preconditions for Takeoff

While in some ways growth of cognitive economics is a success story, it is also a tale of missed opportunities. For more than half a century, the innovative work of Block and Marschak remained largely peripheral in the progression of economic thought. Had Luce, Block, and Marschak joined forces, cognitive economics might have ascended in 1959, instead of taking flight well into the twenty-first century. What's fifty years between friends?

Despite limited direct advances in the decades that followed, a wealth of indirect work quietly laid the groundwork for cognitive economics' eventual ascent. Innovations in behavioral economics, experimental economics, and applied economics consistently underscored the necessity of understanding erroneous decisions. The emergence of neuroeconomics catalyzed crucial debates over what constitutes valid economic data. It is worth reading Gul and Pesendorfer's "In Defense of Mindless Economics" and Camerer's "In Defense of Mindful Economics" to see the arguments made on either side of this case.

Cognitive economics as I see it occupies something of an intermediate space. It pursues a structured process of data enrichment. The starting point is a limitation of Gul and Pesendorfer's argument in defense of mindlessness economics. They make a purportedly Samuelsonian insistence on behavioral data as the sole legitimate economic dataset. What matters is what we choose, not what we think. There is one small catch in this argument. It is based on a currently undefined and possibly undefinable conception of *choice*.

The deep problem as I see it is that there is absolutely no scientific basis for referring to some outputs of human activity as chosen and others not. For an outside observer trying to understand what they are observing, choice data includes any observable that is not entirely predictable. Our intuitive differentiation between chosen and non-chosen activities often stems from subjective experiences of conscious effort and deliberation. However, this internal perception is not readily amenable to scientific investigation with current methodologies. Capturing this aspect effectively demands a quantum leap in our measurement techniques. I can see no basis for saying that our bodily functions are excluded from choice behavior. Pulse fluctuates and as such appears to be chosen from a feasible set subject to constraints. It does not matter from a scientific point of view that we are not consciously aware of this choice. Those in search of a substantial conceptual challenge might ponder this: What would the ideal measurements look like that could breathe scientific life into models of consciousness and conscious choice? At present, this remains a formidable challenge, one that continues to elude definitive answers. I have an instinct that there is a valid path forward in which the subjective experience of consciousness can drive learning forward at a faster rate, but that is a subject for another day. I would be happy to interact with any who are interested in pursuing this further.

My current research focus is different. Well before we have operational measures that allow us to differentiate between conscious and unconscious choice, we will incorporate biological factors into our models of choice. Caplin et al. (2010) incorporate neural data into the estimation of *reward prediction error* in models of *reinforcement learning*, building on Schultz et al. (1997) and Bayer and Glimcher (2005). The constraint here is essentially practical. Measurement technologies are not yet available to test important theories of neural signals on a real-time basis. Application of such data is currently hampered by the need for significant advances in real-time measurement technologies. The bottom line is that the endeavor to categorize human activities as either chosen or mechanical, without considering the limitations and potential of measurement technology, is likely to falter. Practical constraints are often temporary, and should not be mistaken for fundamental limitations.

8.4 Stage 3: Takeoff to Sustained Growth

The recent period has witnessed a notable acceleration in cognitive economic research, driven by interrelated advances in modeling and measurement. As economists increasingly integrate concepts from cognitive psychology, the field has seen profound developments. Almås et al. (2024) summarize many important developments, particularly in terms of survey methodology, that dig further into beliefs and preferences, as well as the structure of decision-making within the family. For present purposes, possibly the most important set of innovations in measurement relates to costs of learning. Pride of place in the theoretical literature underlying these innovations are the theory of search pioneered by Stigler (1961), and rational inattention theory, pioneered by Sims (2003), with a comprehensive recent survey by Maćkowiak et al. (2021). This line of modeling offers a rigorously disciplined approach to defining the limits of rationality and understanding the impact of cognitive limitations on decision-making errors.

In this book, I emphasize the implications theories of cognitively constrained decision-making hold for ideal data collection. The state-dependent stochastic choice data (SDSC) that I introduce in earlier chapters is perfectly suited to rational inattention theory, much like data on choice from all budget sets ideally serves utility maximization theories (Caplin & Dean, 2015). SDSC not only makes rational inattention theory testable but also facilitates distinguishing between utility, learning, and the costs associated with learning. When standard models based on these constructs fail, SDSC provides invaluable insights into the nature of these failures and points to additional factors that may require modeling.

Historically, the introduction of SDSC into cognitive economic research directly responds to Block and Marschak's early calls for new basic observations to delineate desires from cognitive constraints. SDSC serves standard models of cognition in a manner analogous to how traditional choice data has supported classical utility maximization. In the psychometric tradition, SDSC reveals perceptual limits as individuals strive to answer perceptual questions accurately. Extending this to cognitive economics, the literature on SDSC employs revealed preference tools to assess both the inputs (attentional effort) and outputs (subjective improvement in decision quality) of cognitive processes, without presuming prior knowledge of utility.

SDSC's value lies in its ability to differentiate what decision-makers know from what they prefer. By definition, it records choices where payoffs depend on facts that may be obscure to the decision-maker but apparent to the observer. This clarity enables the identification of errors in understanding. For instance, Kőszegi and Rabin (2008) utilize the strict monotonicity in preferences over wealth to demonstrate that bets placed on states of the world can elucidate beliefs and, consequently, cognitive constraints. Cognitive economic models further reveal that much can be discerned from SDSC even when utilities must be inferred by the econometrician, as in classical revealed preference theory.

Technically, classical revealed preference theory hinges on the principle that feasible but unchosen options cannot be preferred to the chosen options. This principle reappears when using SDSC to infer utilities and cognitive constraints. The challenge lies in identifying the right feasible but unchosen options for comparison. The first application of this counterfactual logic appeared in Caplin and Martin (2015), which introduced a "no improving action switch" condition to test for improving switches based on the decision-maker's beliefs. By revealed preference principles, no such wholesale action switches can enhance utility. To align with rational inattention theory, additional conditions are necessary to ensure that no cycle of data among decision problems can improve expected utility. This defines the "no improving attention cycles" condition of Caplin and Dean (2015).

Given that SDSC is a sophisticated generalization of psychometric data, it is frequently gathered in experimental settings designed to pinpoint features of the cost function, such as the experiments by Dean and Neligh (2023). These carefully designed experiments are crucial for evaluating communication quality, work skills, AI performance, and teaching methodologies across various chapters of this book. The versatility of SDSC extends beyond rationality-based theories to also enrich behavioral economic theories, particularly in measuring miscalibration and understanding procedural decision-making. The potential applications of SDSC are vast and continue to expand, promising further insights and advancements in cognitive economics.

There is important early stage work on using SDSC to test models of strategic interaction (e.g. de Clippel & Rozen, 2020). There is much to be done in this regard. There are also directly cognitive approaches to game theory, in particular the level-K model of Nagel (1995) and the cognitive hierarchy model of Camerer et al. (2004) . These too call for innovative measurement. For example, Agranov et al. (2015) introduce data on how choices evolve over time to estimate levels of strategic sophistication.

Just as economists have increasingly developed operational models of costly learning, so have cognitive psychologists in developing models of resource rationality (Griffiths et al., 2015). The driver is very different. The cognitive psychologists are trying to bring order to the large literature on heuristics and biases (Gigerenzer & Gaissmaier, 2011). They are seeking to better understand why and when certain heuristics are chosen over others. The theory of resource rationality is designed to do just this, using cost-benefit trade-offs between the algorithms. This convergent evolution across disciplines strongly suggests interdisciplinary convergence.

Another vibrant line of cognitive economic research, exemplified by Woodford (2020), models apparent deviations from optimal behavior as resulting from optimally imprecise internal representations of available options. This line of research investigates whether behaviors often labeled as irrational have cognitive foundations. For instance, Rabin and Thaler (2001) highlight a phenomenon where people exhibit risk aversion over very low stakes, which traditional economic models fail to explain. Some explanations are behavioral: e.g. narrow framing. But there are now cognitive models suggesting that reasonable errors in translating objective outcomes to subjective perceptions might explain this behavior (Khaw et al., 2021).

As models of the economic impact of cognitive constraints develop, researchers are running into ever more links with the psychological tradition. Economists, inspired by psychologists, are applying principles of efficient neural coding. For example, these principles explain why the valuation of an option depends on available alternatives rather than on its intrinsic value. Woodford (2012) demonstrates how efficient coding and Bayesian decoding together offer a concise model of such context effects. Efficient coding principles also capture how subjective valuations' sensitivity to objective situations depends on the distribution of values used in experiments (Frydman & Jin, 2022). These models suggest that many seemingly irrational behaviors may be optimal when considering cognitive resource limits.

Various other branches of cognitive economic research, particularly those focusing on attention, are thriving. One such area is salience theory. Lichtenstein and Slovic (1971) show how different gambling environments shift focus between prize amounts and associated probabilities. Rubinstein (1988) introduce another interaction between attention and utility, proposing that attention is drawn to dissimilarities. Kőszegi and Szeidl (2013) model how attention is drawn to the most differing attributes, while Bordalo et al. (2022) develop a theory of salience-based utility reweighting.

For the most part, these research lines place little emphasis on innovations in measurement. However, as emphasized throughout this book, significant advances in cognitive economics often stem from diversifying data forms. Examples include survey measurements, process measurements, cognitive instruments, and SDSC. Future innovations will likely explore areas like memory limitations and beliefs about algorithmic advisors. As new data forms develop driven by advancements in measurement technology, computational abilities, and biological sciences, we will finally realize the importance of Block and Marschak's prescient call for new basic observations.

Another key area of ongoing cognitive economic research is decision time. Decision time often correlates with the level of attention given to a problem, with longer times indicating better task performance. Psychometricians have long studied the relationship between decision time and quality, particularly through the drift-diffusion model (Evans et al., 2019; Ratcliff, 1978), which has recently been incorporated into economics (Alos-Ferrer et al., 2021; Fehr et al., 2011).

Indicators of search behavior are also being used to understand choice behavior better. Johnson et al. (2002) pioneered the mouselab interface that records whether decision-makers choose to observe certain items, challenging models with strategically sophisticated players. Gabaix et al. (2006) used similar interfaces to model search in complex environments, while Reutskaja et al. (2011) employed eye-tracking to study item consideration before making choices.

Limits on memory represent another active research area in cognitive economics. Akerlof and Yellen (1985) showed limited recall of past unemployment experiences, leading to early literature on salience effects and memory fading (Topel, 1990). Malmendier and Nagel (2011, 2016) demonstrated that beliefs about the future are strongly influenced by past experiences, particularly those from formative years. da Silveira et al. (2020) applied efficient coding principles to derive optimal memory patterns, while Bordalo et al. (2017) modeled salience effects in memory. Malmendier and Wachter (2021) explored how current events trigger memories of analogous past periods, influencing perceptions and decisions.

8.5 Stage 4: The Drive to Maturity

We are currently at the start of Stage 4 in the development of cognitive economics: the drive to maturity. What is setting this in motion is a growing sense of community. Some of the important events marking takeoff of the larger discipline are associated with the Sloan-NOMIS Program on the Cognitive Foundations of Economic Behavior. Ernst Fehr, Michael Woodford and I lead this Program, which is charged with lowering barriers to communication across fields and the disciplines through a series of conferences, workshops, and summer schools. The summer schools in 2018, 2019, and 2022 were each attended by some thirty young scholars who were exposed to the converging frontiers of economic and psychological research.

There are many complementary developments that point to convergence between economics and cognitive psychology. The annual Bounded Rationality in Choice Conferences, led by Paola Manzini and Marco Mariotti, date back to 2014. They are strongly linked with progress in cognitive economics. There is also a modern re-birth of interest in information economics, sparked in large part by the work of Kamenica and Gentzkow (2011) on Bayesian persuasion. Behavioral economics is also taking an increasingly cognitive turn (e.g. Gabaix, 2014), and there is an increased focus on data innovation both in survey design and in the experimental laboratory (Enke & Graeber, 2023). Cognitive psychologists hold highly complementary annual Workshops on Cognitive Effort.

Despite these clear signs of convergence, challenges remain. Specifically, very few research teams have formed across disciplinary boundaries, and those that have done so operate on a relatively small scale. This is largely due to the funding model in social scientific research, which is both limited and archaic. There is much unrealized potential for merging of interests and formation of research teams that cross disciplinary boundaries, but this will be liberated only with a change in funding priorities. Chapter 10 makes proposals in this regard.

References

Adams, A., Cherchye, L., De Rock, B., & Verriest, E. (2014). Consume now or later? Time inconsistency, collective choice, and revealed preference. *American Economic Review, 104*(12), 4147–4183.

Agranov, M., Caplin, A., & Tergiman, C. (2015). Naive play and the process of choice in guessing games. *Journal of the Economic Science Association, 1*, 146–157.

Akerlof, G. A., & Yellen, J. L. (1985). Unemployment through the filter of memory. *The Quarterly Journal of Economics, 100*(3), 747–773.

Almås, I., Attanasio, O., & Jervis, P. (2024). Economics and measurement: New measures to model decision making. *Econometrica, 92*(4), 947–978.

Alos-Ferrer, C., Fehr, E., & Netzer, N. (2021). Time will tell: Recovering preferences when choices are noisy. *Journal of Political Economy, 129*(6), 1828–1877.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron, 47*(1), 129–141.

Block, H. D., & Marschak, J. (1960). Random orderings and stochastic theories of response. In *Economic information, decision, and prediction: Theory and decision library, vol 7-1*. Springer.

Blundell, R., Browning, M., & Crawford, I. (2008). Best nonparametric bounds on demand responses. *Econometrica, 76*(6), 1227–1262.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2017). Memory, attention, and choice. *The Quarterly Journal of Economics, 131*, 1753–1794.

Bordalo, P., Gennaioli, N. and Shleifer, A. (2022). Salience. Annual Review of Economics, *14*(1), pp.521–544.

Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics, 119*(3), 861–898.

Caplin, A., & Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review, 105*(7), 2183–2203.

Caplin, A., Dean, M., Glimcher, P. W., & Rutledge, R. B. (2010). Measuring beliefs and rewards: A neuroeconomic approach. *Quarterly Journal of Economics, 125*(3), 923–960.

Caplin, A., & Martin, D. (2015). A testable theory of imperfect perception. *Economic Journal, 125*(582), 184–202.

da Silveira, R. A., Sung, Y., & Woodford, M. (2020). *Optimally imprecise memory and biased forecasts* (National Bureau of Economic Research Working Paper 28075).

Dean, M., & Neligh, N. (2023). Experimental tests of rational inattention. *Journal of Political Economy, 131*(12), 3415–3461.

de Clippel, G., & Rozen, K. (2020). *Communication, perception, and strategic obfuscation* (Working Paper).

Enke, B., & Graeber, T. (2023). Cognitive uncertainty. *The Quarterly Journal of Economics, 138*(4), 2021–2067.

Enke, B., Graeber, T., & Oprea, R. (2023). *Complexity and time* (NBER Working Paper Series, No. 31047).

Enke, B., & Zimmermann, F. (2019). Correlation neglect in belief formation. *The Review of Economic Studies, 86*(1), 313–332.

Evans, N. J., Holmes, W. R., & Trueblood, J. S. (2019). Response-time data provide critical constraints on dynamic models of multi-alternative, multi-attribute choice. *Psychonomic Bulletin and Review, 26*, 901–933.

Fehr, E., & Rangel, A. (2011). Neuroeconomic foundations of economic choice—Recent advances. *Journal of Economic Perspectives, 25*(4), 3–30.

Frydman, C. and Jin, L.J. (2022). Efficient coding and risky choice. *The Quarterly Journal of Economics*, *137*(1), pp.161–213.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, *129*(4), pp.1661–1710.

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review, 96*(4), 1043–1068.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451–482.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science, 7*(2), 217–229.

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory, 104*, 16–47.

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review, 101*(6), 2590–2615.

Khaw, M.W., Li, Z. and Woodford, M. (2021). Cognitive imprecision and small-stakes risk aversion. *The review of economic studies*, *88*(4), pp.1979–2013.

KőSzegi, B., & Rabin, M. (2008). Revealed mistakes and revealed preferences. In A. Caplin & A. Schotter (Eds.), *The foundations of positive and normative economics: A handbook* (pp. 193–209). Oxford University Press.

Kőszegi, B., & Szeidl, A. (2013). A model of focusing in economic choice. *The Quarterly Journal of Economics, 128*(1), 53–104.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of experimental psychology, 89*(1), 46.

Luce, R. D. (1958). A probabilistic theory of utility. *Econometrica, 26*, 193–224.

Maćkowiak, B., Matějka, F., & Wiederholt, M. (2021). Rational inattention: A review. *Journal of Economic Literature, 61*(1), 226–273.

Malmendier, U., & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics, 126*(1), 373–416.

Malmendier, U., & Nagel, S. (2016). Learning from inflation experiences. *The Quarterly Journal of Economics, 131*(1), 53–87.

Malmendier, U., & Wachter, J. A. (2021). *Memory of past experiences and economic decisions*. Available at SSRN 4013583.

Manzini, P., & Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review, 97*(5), 1824–1839.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American economic review*, *85*(5), pp.1313–1326.

Rabin, M. and Thaler, R.H. (2001). Anomalies: risk aversion. *Journal of Economic perspectives*, *15*(1), pp.219–232.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review*, *85*(2), p.59.

Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review, 101*(2), 900–926.

Rostow, W.W., (1959). The stages of economic growth. *The economic history review*, *12*(1), pp.1–16.

Rubinstein, A. (1988). Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?). *Journal of Economic Theory, 46*(1), 145–153.

Samuelson, P. A. (1938). A note on the pure theory of consumer's choice. *Economica, 5*, 61–71.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593–1599.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics, 50*, 665–690.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy, 69*(3), 213–225.

Topel, R. (1990). Specific capital and unemployment: Measuring the costs and consequences of job loss. In *Carnegie-Rochester conference series on public policy* (Vol. 33, pp. 181–214). North-Holland.

Weber, E. H. (1834). *On the Tactile senses* (H. E. Ross & D. J. Murray, Trans. and Eds.). Experimental Psychology Series.

Woodford, M. (2012). Prospect theory as efficient perceptual distortion. *American Economic Review, 102*, 41–46.

Woodford, M. (2020). Modeling imprecision in perception, valuation, and choice. *Annual Review of Economics, 12*, 579–601.

CHAPTER 9

Next Steps in Research, Business, and Policy

This chapter outlines next steps in developing cognitive economic research methods and applying them in the field. I open the chapter by outlining next steps in the research agendas outlined in Chapters 5, 6, and 7. I focus in particular on research related to the cognitive economy: how to improve human-AI collaborative decision-making; how to better inform workers and students about the opportunities and risks associated with the rise of AI; and how to teach the skills that will matter in the cognitive economy. I include in this some personal conjecture on how best to incorporate the rise of AI in general, and large language models in particular, into teaching programs for students of all ages. Perhaps not surprisingly in light of my professional and personal identity, I do not suggest that others adopt these ideas until they are put to the test in the manner I propose.

In addition to sketching out important forward-looking research projects, my goal in this chapter is to show that cognitive economics is very much ready for business. I close the chapter by outlining a few of the many business and policy advances that rest on development and application of cognitive economic methods.

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_9

9.1 Research Area 1: Human-AI Decision-Making Pipelines

The first open research area concerns designing the full human-AI decision-making pipeline in important use cases, such as medical decisions, credit decisions, hiring decisions, etc. At present, most research on human-AI collaboration concerns how they are integrated into final decisions. Is it better to have a human-in-the-loop to interrupt otherwise poor algorithmic decisions, a machine-in-the-loop to step in when humans are messing up, or is there some better way to coordinate human and machine in decision-making?

Important as is this final decision stage, it should not be seen in isolation. There are many other interactions between human and AI in the earlier stages of the decision-making pipeline. As noted in Chapter 5 this typically begins with human labeling and annotation of instances. I outline therein research on the labeling industry, and point to the long and fruitful path forward. Also noted in the chapter is the fact that data scientists are also human, and that their methods can potentially be improved upon using insights from cognitive economics.

Putting it together, what is really missing are holistic studies of the human-AI decision-making pipeline. We need to know how to set this pipeline up to take advantage of the massive potential to raise decision quality while avoiding a *cascade of biases* that might otherwise emerge from the ill-coordinated Rube Goldberg kluge that defines most human-AI decision-making protocols today. In many cases, there are grounds for worry about data cascades in which mistaken labels cause AIs to learn poorly, give flawed advice, and then pass this on to biased humans who do not really understand what the AIs mean. This is particularly a concern in terms of biases that show up in arrests, credit offers, job offers, etc. In the worst of cases, biases may get amplified as low quality and biased labels, poorly chosen loss functions, and inappropriate and further biased human-AI interactions do more harm than good.

Cognitive economic research methods are now in place to allow holistic analysis of the many levels at which human and algorithmic resources are integrated into the decision-making pipeline. Research in this area is conceptually easy. All one needs to do is to appropriately measure the quality of final decisions as one varies methods at all stages of the pipeline. We need rich experimental architectures in which changes at all stages of the pipeline are explored to drive an understanding of how they contribute separately and together to decision quality.

While conceptual issues are straight forward, implementation is anything but. The required research is profoundly multi-modal and interdisciplinary. It involves economic modeling, experimental design, psychological expertise, and machine learning, along with subject matter experts in the fields of application, which might, e.g. be medical or related to job hiring. This need for interdisciplinary collaboration and for commercial partnerships is already evident in the research outlined in Chapter 5 on labeling of medical images. This research relies not only on collaboration between economic theorists, cognitive psychologists, data scientists, and experimentalists, but also on collaboration with Centaur labs and medical experts from Vanderbilt University Medical Centre. Supporting such collaborations on a larger scale will not come cheap.

9.2 Research Area 2: Is Ignorance About AI Harming Careers?

The ongoing cognitive revolution will cause massive turnover in the skills that are important for earning money over the course of a career. We need to ramp up research on opportunities and threats associated with the cognitive economy as perceived by workers of all ages and students and the preparatory measures they are taking in response to these perceptions. We also need to assess the realism of these assessments, and find methods for closing identified gaps with reality.

In an ideal case, workers of all ages will know which skills would provide them with the flexibility that will be needed. Even if not, information would be assertively conveyed to make us all as aware as we should be about our educational, training, and retraining options and where they might lead. We would know what skills to hone to improve our future prospects. We would be as aware as humanly possible of what the future might hold in terms of changing job opportunities. We would still be surprised, but we would have plans in place to retool. We would face our essentially limitless options in a thoroughly informed manner. Not all would be well with the world, but much would be better.

We need to research not the ideal, but the real. We need to understand how aware workers *actually are* of the opportunities and threats that AI brings. The same applies to students, who are currently choosing how to prepare for the workforce. If they are unaware of how the labor market is changing, then they may be ill-prepared for what the future will bring. This may foment later despair as workers' visions of a stable future are dashed. Identifying individuals and groups who are unaware of the likely impact of AI on their earnings if they do not upgrade skills will help identify possible sources of future economic and social instability, and allow resources for conveying information on the skills of the future and correspondingly retraining workers to be directed where they are most needed.

In addition to looking at issues from the worker side, researchers need to take a complementary look at employers' views of the impact of AI on their need for skills, decision-making, and otherwise, as well as how they believe it will impact the organization of the labor force. Developing complementary surveys of workers and firms will make it possible to identify mismatches between worker and firm expectations. These may be signals of impending trouble, with workers being relatively unaware of the fragility of their situations at work. In that way, we may be able to produce technological early-warning systems for mismatches in beliefs about the future warranting deeper investigation and ameliorative action. As always it is important to identify methods of conveying information to aid educational and career decisions.

What makes this an opportune time to ramp up research effort is that key research methods and research resources are now in place. With regard to methods, Chapter 6 closes by outlining early stage research on workers' perceptions of future job prospects, while Chapter 7 closes by outlining research on students' perceptions of returns to education, and how to correct any illusions they are found to have. With regard to resources, the Copenhagen Life Panel (CLP) introduced in Chapter 4 is uniquely valuable. It is conducted annually with a random sample (at least as near as possible to random) from the Danish population registries. The survey gathers detailed responses from more than 10,000 respondents concerning the beliefs about future employment and earnings. The key to its value is that it asks questions about possible job transitions and their impact on future earnings. Specifically it elicits subjective probabilities of being laid off, staying in the current job or quitting. It also gathers information on likely time out of work following job separations and about subsequent earnings. The state contingent nature of these beliefs provides insight into how respondents expect to react in scenarios that they are yet to experience. Denmark is a valid place to initiate this line of research given that it is technologically at the forefront and has a particularly flexible and responsive job market.

Importantly, the Danish registry infrastructure allows survey answers to be linked at the individual level to administrative data with labor market outcomes such as earnings, job separations, unemployment, education, industry of employment, wealth, and demographic information. The ongoing and forward-looking nature of CLP gives researchers a unique chance to learn how workers and students adapt to changes in the labor market as the economic transition unfolds. Key questions concern who realizes that AI will change their labor market opportunities, and what they choose to do in the face of possible layoffs. How many would take time out of the work to retrain? How does this willingness to adapt across workers of different ages, education levels, and professional backgrounds? We are currently implementing a first version of this survey of worker experiences with, readiness for, and attitudes toward AI in the workplace. Stay tuned.

The fact that survey data about beliefs can be combined with third-party reported data on subsequent realizations will allow researchers to learn when workers are surprised, and how they respond to such surprises. Comparing labor market beliefs to their realized counterparts will reveal not only how realistic are people's expectations, but also how they react to various kinds of shock.

An ideal research protocol will provide society with early warnings about mismatches between expectations and likely outcomes, and where these mismatches point to future trouble. It will make it possible to identify who will likely manage to transition successfully and who will fail to adapt appropriately to the cognitive revolution. With enough early warning, it may be possible to intervene to beneficially impact those who appear dangerously unaware of the high risks they face as the economy transitions.

9.3 Research Area 3: Teaching for the Cognitive Economy

Cognitive economic research offers many promising paths forward in relation both to identifying and to teaching the skills that will matter for careers in the cognitive economy. Key questions are how to teach newly valuable skills to those entering the workforce and how to retrain older workers whose skills have been devalued.

The first research beach-head is outlined in Chapter 6, which suggests that various forms of self-knowledge, such as calibration, as well as various forms of algorithmic awareness, may be of value in work with an AI. The obvious next step is to find out how general these skills are and how teachable they are. There are 100 other such skills waiting to be explored.

In addition to skills that help in a given job, Chapter 6 suggests also career skills that may be important, particularly those associated with making moves up the job ladder by finding better and better jobs. As we research how search skills might matter in deciding when and where to apply for jobs, we also need to know the extent to which these are generalizable and teachable. Education and training for the cognitive economy may have little in common with the education of today. What is important when ChatGPT writes resumes is anybody's guess. However this turns out, it is important to know more as soon as possible.

As promised, I now provide an opinion that derives from my role as teacher more than as researcher. All of us who want students to participate in creative writing tasks (I am one) face a new challenge in the age of AI: how to deal with the availability of large language models (LLMs). The Luddites would like to ban them, much as Socrates disdained the written word as not expressing *real* thought. I am not sympathetic, in case my choice of analogy left you uncertain. So if banning the use of LLMs is out of the question, how do we ensure that our students do not get their answers without having the slightest ability in self expression let alone anything interesting to say?

My proposal is to reward both what is written and the interactive process that generated it. Henceforth we should offer grades both for the final project write-up and *for an annotated and verified write-up of the process of production*. The project write-up would be graded as now, based purely on its quality, with no regard to how it was generated. This quality would account for a certain proportion of the final grade: say 50% to keep it simple. The rest of the grade would be determined entirely differently. It would be based on a time-stamped write-up of how precisely the answer was created, including all interactions with LLMs. This may in fact be the more important skill to hone in the long run, more than the project itself.

There are reasonable counter-arguments. For example, some might say this has the down-side that our students will focus as much on process as on personal investigation of the subject matter. I agree with that this would be possible, but conjecture that even in this regard we will learn more in the process. After all, LLMs have a lot to teach us about essentially everything, provided students learn how to check, expand, and adapt to their input. Indeed part of an ideal education would be teaching this process in a staged process with a deliberately misinformed LLM. More importantly it is the higher level process skill that is nowadays more important than any particular project. In future the key skill for essentially all tasks will be the ability to interact effectively with all the AI agents that are available to end up with an appropriate personalized synthesis. What an exciting new age of creativity that will be. New grading schemes need to be tested with some urgency if we are to get ahead of this future. To implement would require the teachers themselves to be adept with LLMs, a far stretch. If it turns out as I expect, we will need to change teacher training correspondingly. A small price for teachers to pay for students to earn a living in the cognitive economy.

9.4 Business and Policy Applications

Many should be interested in cognitive economic research making rapid progress, including policymakers and business leaders. There are many cognitive economic research protocols and findings that might be important in business and policy settings. I outline a few obvious cases that are of particular personal interest. There are many others that readers may understand better than I. Note that I will not differentiate sharply between business and policy openings. I am only outlining business opportunities whose rationale is to create important social benefits. Likewise I am only interested in policymakers who are focused on social welfare. Whether policymakers or businesses will make more valuable contributions is likely to differ on a case-by-case basis.

- Chapter 2 introduces cognitive household finance which is replete with opportunities for social improvement. Here are two areas of particular interest.

    1. It is crystal clear that savings and portfolio choice depend on cognitive factors. Chapter 2 highlights the importance of *financial literacy*, *financial education*, *self control*, and *financial planning*. At present there are a number of websites that offer help in this regard, as well as a number of commercial players. We know little as yet about how effective these

measures are. Clearly further research in these areas will be of massive importance to financial advisers and financial institutions. Just last week I passed my local Chase branch, and it invited me in to discuss *The Power of Planning*. The work of Ameriks et al. (2003) suggests there may be something to this. Better late than never I guess. Even the speed of business is not all it is cracked up to be. I suspect Chase that at this stage, Chase knows no more about this than we did back then: it is more slogan than science at this stage. It would be good to gather more granular scientific evidence and thereby provide more value to customers.

2. Testing and protecting against cognitive decline are areas of great importance both for the social security administration and for private asset managers. Yet there is as yet little to no research on how to better protect either social security recipients or elderly asset holders against their own future decline and the financial dangers that it brings in train. There is plenty of room for asset managers to develop helpful services in this area in terms of flagging transactions and following up as appropriate with clients and indicated family members. There is also room for financial innovation. An area that I have researched in some depth is creation of new solutions to allow elderly homeowners to remain in their homes rather than move to care facilities. The home is the obvious asset to use to pay for care for those who wish to remain in the home in their later years. Equity sharing markets, which are beginning to develop along lines sketched out by Caplin et al. (1997), have much potential in this regard as debt-free alternatives to reverse mortgages.

- Chapter 3 concerns measuring and reducing the damaging effects of complex communication:

   1. It indicates need to conduct reforms of the justice system aimed at reducing high error rates. It is hard to understand why legal scholars and practitioners have not developed methods to measure yet alone reduce errors. What exactly is holding them back? Whatever it is, it is past time for them to step up to the plate.

2. Policymakers interested in consumer protection can readily adapt the experimental methods introduced in Chapter 3 to test the effects of possible disclosure regulations in reducing decision-making mistakes. This is essentially shovel-ready policy research, although of course the design will differ in details on a case-by-case basis. It is well timed, since there is huge interest already in reducing the burden that complex policies impose on households, as witnessed in Executive Order 14058 (2021) on rebuilding trust in government.
3. In addition to well-formulated regulatory solutions, which I believe to be needed, there is clear room for private businesses to offer certification services by scientifically testing for clarity of communication using experimental methods in combination with AI. The business idea is to put in place a private *simplicity certification service*, much like a Cognitive Economic Underwriters' Laboratory. These methods would also be of interest to businesses such as Apple that wish to highlight privacy protection and other forms of responsiveness to consumer concerns.

I could go on, but you get the point. There are countless other ideas of equal validity that may, for all I know, be closer to implementation. This is the tip of a very large iceberg. I look forward to a day when business and policy clusters form around leading institutions in cognitive economics, as in the biological and data sciences. In the next and final chapter, I lay out a path to speeding arrival of this day, and indicate how each of us can play a part.

Reference

Ameriks, J., Caplin, A. and Leahy, J. (2003). Wealth accumulation and the propensity to plan. *The Quarterly Journal of Economics*, *118*(3), pp.1007–1047.

Caplin, A., Chan, S., Freeman, C., & Tracy, J. (1997). *Housing partnerships*. Cambridge, MA: MIT Press.

CHAPTER 10

Accelerating Cognitive Economics: Why Now, and How?

I want you to share my sense of urgency. If we continue with business as usual, cognitive economics will make only slow and incremental progress. Those of us who are open to collaboration will continue to push forward. We will also continue to hold interdisciplinary conferences to advance mutual understanding. But this will not change the outdated structure of the social scientific research enterprise. Support for the critical team-building efforts needed for cognitive economics will still be missing. Publication lags in economics, which can see more than 5 years pass between execution and publication, will continue to be absurd. The connection with policy will continue to be indirect and subject to even further delays. We will also continue to rely on an archaic and restricted funding model that has done little to support the forms of team research cognitive economics requires.

What is needed for more rapid and durable progress in cognitive economics is institutional change that supports the development of effective research teams. In this final chapter I lay out a future in which a new institution, which I shall call the Accelerator for Cognitive Economics, plays the ideal role in developing, applying, teaching, and promoting cognitive economics. I will use the present tense and write as if the Accelerator exists. I would like all of us to imagine this ideal future in a visceral enough manner to vividly experience what we are currently missing. This may provide the motivation for the maximum number of readers and other interested parties to join together in making the required changes.

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9_10

I close the chapter and with it the book by proposing steps that key parties can take to contribute to the growth of cognitive economics in the short run, before there has been the opportunity for institutional reform. There are specific call-outs for fellow researchers, policymakers, business leaders, entrepreneurs, students, teachers, and socially minded members of the general public. Working together we have the opportunity to accelerate this important new science, to the benefit of broader society.

10.1 An Accelerator for Cognitive Economics

A central role of the Accelerator is to support research teams in cognitive economics. Literally no one can effectively learn the skills that are needed for effective research in cognitive economics. Ideal research teams will have to include not only economists and cognitive psychologists, but also other social scientists, data scientists, and specialists in application. Collaborative work is crucial not only in advancing theoretical knowledge but also in providing new forms of data that can inform better policy decisions and business practices. As researchers push the boundaries of what is known, they will also need to engage with practitioners in business and policy to ensure that their findings are applicable and actionable.

The Accelerator treats newly formed research teams as start-up operations. There is a commitment on the part of the researchers and funders to ensure that the results are written up, published, and made as broadly available as possible. The research teams identify clients with whom they can agree valid progress metrics. Judgment of success hinges not only on the opinions of other research leaders and academic peers, but also on representatives of public policy, community leaders, and the commercial world. There is funding for pre-doctoral and post-doctoral students to join collaborative research projects. This provides essential training for younger scholars setting out on interdisciplinary research. Participating junior researchers will play an essential role in eroding artificial barriers between current social scientific disciplines.

The governance of the Accelerator includes a board comprising leading figures from academia, industry, and government to provide strategic direction and oversight. The Accelerator hosts regular conferences and symposiums to share updates, discuss challenges, and refine strategies. It has a significant educational mission, designing curricula, hosting pre-doctoral and post-doctoral students, and linking them up with other educational and research institutions. It operates through four core divisions, each focused on a critical aspect of cognitive economics.

- The Scientific Division focuses on advancing cognitive economics as a science. It is highly connected with research departments in universities, policy think tanks, and businesses. It supports the development of new cognitive economic models and their applications. Critically, it also supports the forms of team-building required to implement these models in practice. There are new forums for researchers to initiate cognitive economic research projects of appropriate scale and to identify potential collaborators. The Division promotes new methods of grant funding that are on a larger scale than currently available for social scientific advance, and that are staged. An advantage of this funding model is that it creates far closer ties between the world of research and that of application. Researchers get real-time feedback on which ideas are important to pursue now, which can be pursued later, and which are dead ends.
- The Research and Development Division includes a Business and Innovation Hub and a Policy Lab. The Business and Innovation Hub ensures that research with practical applications aligns with market needs and policy developments. A Business Incubator supports startups and organizations by translating research insights into practical applications, pilot programs, and prototypes. The Policy Lab provides data-driven recommendations for social programs, education campaigns, and reforms. The Research and Development Division proactively reaches out to businesses and policymakers who face challenges that may be overcome with cognitive economic research. In addition to hosting forums, the Research and Development Division is responsible for introducing scientific findings to policy institutions, businesses, educational institutions, etc.
- The Education and Training Division offers educational programs ranging from executive courses to full academic degrees, utilizing AI-driven platforms for personalized learning experiences. It provides a curated list of resources for deeper learning, including contact information for relevant organizations and thought leaders. It hosts an educational platform that offers accessible courses, workshops, and conferences with cutting-edge insights. It also develops educational

programs and outreach initiatives to prepare current and future generations for roles in this evolving field. It hosts pre-doctoral as well as post-doctoral students. It hosts dedicated meetings for students and teachers interested in advancing the field of cognitive economics. Educators have the chance to discuss innovative teaching methods, curriculum development, and potential research projects that can contribute to the growth of this discipline. By engaging with students and teachers in these discussions, researchers are more rapidly able to drive the field of cognitive economics and social science as a whole forward.
- Publication and Outreach Division. This division is charged with involving as wide a community as possible in social scientific advance. The publication arm of the Accelerator creates new outlets for ideas and approaches to become part of the academic and popular discourse. This includes programmatic books by leading researchers, as well the conference volumes that have made data science the most successful new discipline in many decades. Outreach is also important. Social scientific research needs to hit a chord with a far larger audience to grow effectively. Having experts in modern methods of communication will be critical. The academy is very far from an appropriate role model in this regard.

The issue of publication is particularly intricate and crucial. One central reason for writing this book is to fully express ideas on *how to conduct social scientific research*. Academic economics places sharp constraints on publication, with few outlets for *methods and ideas*. It is hard for an outsider even to imagine how absurd and broken the publication process has become in academic economics. There are only 5 publications that really matter, and their criteria for publication are powerfully impacted by the preferences of the editors. Publication can take many years even for those who make it through these filters. Journalists, who might in principle help break the trap by publicizing top research, lack relevant expertise and are subject to their own pressures. As a result they often seek *opinions*, not methods or ideas. This book aims to break these traps. I'm inspired by Bob Shiller's Narrative Economics, which argues that shifts in popular narratives significantly impact macroeconomic dynamics. His book does not meet the top 5 publication criteria any more than this book does, yet it is potentially far more important-just as I hope this book will be. Let 1000 Pivots bloom!

10.2 My Invitation to You

I close by making the promised invitation to cognitive economics. First, I invite you, the readers, to help in building momentum for next steps. I have designed a survey for you to provide feedback. Among other things, this asks for your opinion on the questions that are of most interest to you going forward. I ask specifically about the areas covered in the book but also for suggestions on other priorities. I also look for other feedback on how to accelerate progress in the field, on business or policy ideas, and on community creation. Scanning the QR code above will lead you to the survey of interest and other materials.

SCAN ME

The second invitation is more of a stretch. The Accelerator for Cognitive Economics is just a dream for now. The only way to build toward it is by paving the road with smaller scale initiatives. There need to be meetings of key parties, some virtual and some in-person. Here are some meetings of particular value as building blocks for the discipline.

1. *Researcher Meet-Up with Grant Officers*: Grant officers in government agencies and foundations are the catalysts who can fund the exploration and implementation of cognitive economic theories and technologies. Given that advances in cognitive economics will depend on support and funding from grant officers and grant organizations, we need dedicated meetings of cognitive economic research teams with potential funders. Success of the meeting will be marked by the number of research proposals that result and the innovations and success of these lines of research. Early projects are best designed in a forward-looking manner as demonstration projects.
2. *Researcher Meet-Up with Business Leaders*: There is massive potential for businesses to be the pivotal drivers of change in bringing principles of cognitive economics into our everyday lives. We need to design new financing mechanisms that involve multiple rounds of funding with precise performance metrics and checks on achievement. Dedicated meetings with business leaders and innovators can serve as a catalyst for networking and potential partnerships. These will contribute to the broader adoption and advance of cognitive economic methods in the business world.
3. *Researcher Meet-Up with Policymakers*: Policymakers are crucial in shaping incentives and frameworks that align with cognitive economic principles. They have a unique role in shaping the environment in which the cognitive revolution unfolds. As detailed in Chap. 3, cognitive economics can improve the clarity and effectiveness of public communications, such as regulatory announcements and policy guidelines. To facilitate a comprehensive understanding and foster the implementation of cognitive economic principles in public policy, we need dedicated meetings between researchers and policymakers. This will provide a platform to explore strategies for integrating cognitive economic research into policy frameworks. The goal of these meetings is to create a collaborative environment where

policymakers can learn from each other and from academic experts in the field. Participants will have the opportunity to engage in discussions, share their experiences, and explore collaborative opportunities, all aimed at integrating cognitive economics into effective policymaking.

4. *Researcher Meet-Up with Educators and Students*: The future of cognitive economics relies heavily on the engagement and contribution of students and educators. By engaging deeply with the principles of cognitive economics, students can prepare themselves not just for current technologies but for future innovations. To make this happen will require educators and educational institutions to develop interdisciplinary programs that merge economics and psychology with data science. There should be dedicated meetings for students and teachers interested in advancing the field of cognitive economics. These will serve as a platform to explore the foundations of cognitive economics, discuss its applications, and brainstorm ways to integrate these concepts into academic curricula and research. They will offer students and educators the opportunity to engage with the tangible benefits of cognitive economics and underscore its relevance in various fields. We will explore how cognitive economics can be woven into existing educational frameworks. Educators will have the chance to discuss innovative teaching methods, curriculum development, and potential research projects that can contribute to the growth of this discipline. The meetings can serve as a catalyst for networking and potential collaborations, fostering a community of learners and educators dedicated to advancing cognitive economics. Researchers both contribute their ideas on teaching and absorb lessons on priorities and methods of conveying ideas to those who are less familiar with them.
5. *An Open Forum with Researchers*: There is value in meetings in which engaged members of the public can learn from experts and each other, share their perspectives, and explore ways to advances in cognitive economics. This will provide opportunities for networking and collaboration and empower participants to take an active role in developing and promoting this field. The development and impact of cognitive economics relies not only on experts and policymakers but also on the engagement of socially minded members of the public. Public engagement can spread awareness, making cognitive economics foundational knowledge. This meeting will provide a

> platform to explore the principles of cognitive economics, discuss its societal implications, and brainstorm ways to support and advocate for this emerging field. We will discuss how cognitive economics can address pressing societal issues, such as enhancing public communications, improving decision-making processes, and fostering greater transparency in governance. By understanding these applications, socially minded individuals can become informed advocates for cognitive economics and contribute to its broader acceptance and implementation.

I invite your interest not only in participating in such meetings, but also in taking more significant roles, all the way up to organization and hosting. I look forward to hearing from those who are so interested using the provided QR code to access the interactive portal. Please spread the word and send out the link to others who may be interested in contributing. No time to waste. For my part I will provide updates on progress and provide the results of the questionnaires. I will also provide details on planned meetings, starting with a launch event for the field of Cognitive Economics currently planned for Summer 2026. The event will be streamed live for those interested and recorded for later reference. Please spread the word and send out the link to others interested and willing to contribute to progress.

10.3 A Confession and a Claim

My opening question was whether you would like a child of yours to become a social scientist. It is time for a confession. My own answer is clear-cut. The pros have hugely outweighed the cons for me. It is hard even to convey the deep satisfaction that life as a social scientific researcher has provided. It is not so much a job as a calling. It requires an intricate mix of imagination and grounding in reality. One needs to understand what is, imagine what might be, and search for paths between. It is entirely engrossing, and if a child of yours catches the bug, lucky them and lucky you.

While I'm at it, let me point out how insanely misplaced is the cult of youth in social science. I understand that I have a horse in that race and am no more believable than Steve Ballmer on the iPhone. That doesn't make me wrong. I believe that I have become far more productive as my life experiences have informed my research agenda and as I have participated in ever richer research teams. You knew I would think that. Again, that doesn't make me wrong.

The bottom line is that there is no age bar on engaging with cognitive economics. Many well past the age of fifty will have much of value to contribute, perhaps more than those with little life experience. Come one, come all.

Bibliography

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043–1171). Elsevier.

Adams, A., Cherchye, L., De Rock, B., & Verriest, E. (2014). Consume now or later? Time inconsistency, collective choice, and revealed preference. *American Economic Review, 104*(12), 4147–4183.

Agarwal, N., Moehring, A., Rajpurkar, P., & Salz, T. (2023). *Combining human expertise with artificial intelligence: Experimental evidence from radiology* (No. w31422). National Bureau of Economic Research.

Agranov, M., Caplin, A., & Tergiman, C. (2015). Naive play and the process of choice in guessing games. *Journal of the Economic Science Association, 1*, 146–157.

Akerlof, G. A., & Yellen, J. L. (1985). Unemployment through the filter of memory. *The Quarterly Journal of Economics, 100*(3), 747–773.

Almås, I., Attanasio, O., & Jervis, P. (2024). Economics and measurement: New measures to model decision making. *Econometrica, 92*(4), 947–978.

Almog, D., Gauriot, R., Page, L., & Martin, D. (2024). *AI oversight and human mistakes: Evidence from centre court*. arXiv preprint arXiv:2401.16754.

Alos-Ferrer, C., Fehr, E., & Netzer, N. (2021). Time will tell: Recovering preferences when choices are noisy. *Journal of Political Economy, 129*(6), 1828–1877.

Ameriks, J., Caplin, A., & Leahy, J. (2003). Wealth accumulation and the propensity to plan. *Quarterly Journal of Economics, 118*(3), 1007–1047.

Ameriks, J., Caplin, A., Leahy, J., & Tyler, T. (2007). Measuring self-control problems. *American Economic Review, 97*(3), 966–972.

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9

Ameriks, J., Caplin, A., Lee, M., Shapiro, M. D., & Tonetti, C. (2023). Cognitive decline, limited awareness, imperfect agency, and financial well-being. *American Economic Review: Insights, 5*(1), 125–140.

Andersen, H. Y., & Leth-Petersen, S. (2021). Housing wealth or collateral: How home value shocks drive home equity extraction and spending. *Journal of the European Economic Association, 19*(1), 403–440.

Archsmith, J. E., Heyes, A., Neidell, M. J., & Sampat, B. N. (2021). *The dynamics of inattention in the (Baseball) field* (National Bureau of Economic Research Working Paper 28922).

Arcidiacono, P., Hotz, V. J., Maurel, A., & Romano, T. (2020). Ex ante returns and occupational choice. *Journal of Political Economy, 128*(12), 4475–4522.

Arrieta, G., & Nielsen, K. (2023). *Procedural decision-making in the face of complexity*. Mimeo Caltech.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics, 118*(4), 1279–1333.

Bakos, Y., Marotta-Wurgler, F., & Trossen, D. R. (2014). Does anyone read the fine print? Consumer attention to standard-form contracts. *The Journal of Legal Studies, 43*(1), 1–35.

Bartoš, V., Bauer, M., Chytilová, J., & Matějka, F. (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review, 106*(6), 1437–1475.

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron, 47*(1), 129–141.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Ben-Shahar, O., & Schneider, C. E. (2017). The failure of mandated disclosure. *Russian Journal of Economics and Law, 4*(44), 146–169.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations, 1*(2), 69–186.

Benoliel, U., & Becher, S. I. (2019). The duty to read the unreadable. *Boston College Law Review, 60*, 2255.

Bernheim, B. D., & Garrett, D. M. (2003). The effects of financial education in the workplace: Evidence from a survey of households. *Journal of Public Economics, 87*(7–8), 1487–1519.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review, 94*(4), 991–1013.

Bickel, J. E. (2010). Scoring rules and decision analysis education. *Decision Analysis, 7*(4), 346–357.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, *24*(2), 265–272.

Block, H. D., & Marschak, J. (1959). *Random orderings and stochastic theories of response* (Cowles Foundation for Research in Economics Working Paper 66). Yale University.

Block, H. D., & Marschak, J. (1960). Random orderings and stochastic theories of response. In *Economic information, decision, and prediction: Theory and decision library, vol 7-1*. Springer.

Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook: The cognitive domain*. David McKay.

Blundell, R., Browning, M., & Crawford, I. (2008). Best nonparametric bounds on demand responses. *Econometrica, 76*(6), 1227–1262.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2017). Memory, attention, and choice. *The Quarterly Journal of Economics, 131*, 1753–1794.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Salience. *Annual Review of Economics, 14*, 521–544.

Braddock, C. H., Fihn, S. D., Levinson, W., Jonsen, A. R., & Pearlman, R. A. (1997). How doctors and patients discuss routine clinical decisions: Informed decision making in the outpatient setting. *Journal of General internal Medicine, 12*(6), 339–345.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review, 78*(1), 1–3.

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work* (National Bureau of Economic Research Working Paper 31161).

Camerer, C. (2008). The case for mindful economics. In A. Caplin & A. Schotter (Eds.), *The foundations of positive and normative economics: A handbook* (pp. 193–209). Oxford University Press.

Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics, 119*(3), 861–898.

Cameron, C. M., & Kornhauser, L. A. (2006). Appeals mechanisms, litigant selection, and the structure of judicial hierarchies. *Institutional Games and the US Supreme Court, 173*, 204.

Campbell, J. (2006). Household finance. *Journal of Finance, 61*(4), 1553–1604.

Caplin, A. (2023). *The science of mistakes: Lecture notes on economic data engineering*. World Scientific Press.

Caplin, A. (2025). Data engineering for cognitive economics. *Journal of Economic Literature*.

Caplin, A., Chan, S., Freeman, C., & Tracy, J. (1997). *Housing partnerships*. MIT Press.

Caplin, A., & Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review, 105*(7), 2183–2203.

Caplin, A., Dean, M., Glimcher, P. W., & Rutledge, R. B. (2010). Measuring beliefs and rewards: A neuroeconomic approach. *Quarterly Journal of Economics, 125*(3), 923–960.

Caplin, A., Dean, M., & Leahy, J. (2022). Rationally inattentive behavior: Characterizing and generalizing Shannon entropy. *Journal of Political Economy*, *130*(6), 1676–1715.

Caplin, A., Deming, D. J., Leth-Petersen, S., & Weidmann, B. (2023). *Allocative skill* (National Bureau of Economic Research Working Paper 31674).

Caplin, A., Gomberg, A., & Sadka, J. (2024). *Judging the judges: Indexing complex information reduces injustice* (NBER Working Paper).

Caplin, A., Gregory, V., Leth-Petersen, S., Lee, E., & Saeverud, J. (2023). *Subjective earnings risk* (National Bureau of Economic Research Working Paper 31019).

Caplin, A., & Martin, D. (2015). A testable theory of imperfect perception. *Economic Journal, 125*(582), 184–202.

Caplin, A., Martin, D., & Marx, P. (2022). *Modeling machine learning* (National Bureau of Economic Research Working Paper 30600).

Chemin, M., Chen, D. L., Di Maro, V., Kimalu, P. K., Mokaya, M., & Ramos-Maqueda, M. (2023). *Data science for justice: The short-term effects of a randomized judicial reform in Kenya*. Toulouse School of Economics.

Christian, B. (2021). *The alignment problem: How can machines learn human values?* Atlantic Books.

Conlon, J. J. (2021). Major malfunction: A field experiment correcting undergraduates' beliefs about salaries. *Journal of Human Resources, 56*(3), 922–939.

da Silveira, R. A., Sung, Y., & Woodford, M. (2020). *Optimally imprecise memory and biased forecasts* (National Bureau of Economic Research Working Paper 28075).

Danz, D., Vesterlund, L., & Wilson, A. J. (2022). Belief elicitation and behavioral incentive compatibility. *American Economic Review, 112*(9), 2851–2883.

de Clippel, G., & Rozen, K. (2020). *Communication, perception, and strategic obfuscation* (Working Paper).

De Finetti, B. (1965). Methods for discriminating levels of partial knowledge concerning a test item. *British Journal of Mathematical and Statistical Psychology, 18*(1), 87–123.

Dean, M., & Neligh, N. (2023). Experimental tests of rational inattention. *Journal of Political Economy, 131*(12), 3415–3461.

Delavande, A., & Rohwedder, S. (2008). Eliciting subjective probabilities in internet surveys. *Public Opinion Quarterly, 72*(5), 866–891.

Deming, D. J. (2021). *The growing importance of decision-making on the job* (National Bureau of Economic Research Working Paper 28733).

Deming, D. J. (2022). Four facts about human capital. *Journal of Economic Perspectives, 36*(3), 75–102.

Dominitz, J., & Manski, C. F. (1996). Eliciting student expectations of the returns to schooling. *Journal of Human Resources, 31*, 1–26.

Dominitz, J., & Manski, C. F. (1997). Using expectations data to study subjective income expectations. *Journal of the American Statistical Association, 92*, 855–867.

Enke, B., & Graeber, T. (2023). Cognitive uncertainty. *The Quarterly Journal of Economics, 138*(4), 2021–2067.

Enke, B., Graeber, T., & Oprea, R. (2023). *Complexity and time* (NBER Working Paper Series, No. 31047).

Enke, B., & Zimmermann, F. (2019). Correlation neglect in belief formation. *The Review of Economic Studies, 86*(1), 313–332.

Epper, T., Fehr, E., Fehr-Duda, H., Kreiner, C. T., Lassen, D. D., Leth-Petersen, S., & Rasmussen, G. N. (2020). Time discounting and wealth inequality. *American Economic Review, 110*(4), 1177–1205.

Evans, N. J., Holmes, W. R., & Trueblood, J. S. (2019). Response-time data provide critical constraints on dynamic models of multi-alternative, multi-attribute choice. *Psychonomic Bulletin and Review, 26*, 901–933.

Executive Order. (2021, December 13). *Transforming federal customer experience and service delivery to rebuild trust in government*.

Fehr, E., & Charness, G. (2025). Social preferences: Fundamental characteristics and economic consequences. *Journal of Economic Literature*

Fehr, E., & Rangel, A. (2011). Neuroeconomic foundations of economic choice—Recent advances. *Journal of Economic Perspectives, 25*(4), 3–30.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *Quarterly Journal of Economics, 129*(4), 1661–1710.

Gabaix, X., & Laibson, D. (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics, 121*(2), 505–540.

Gabaix, X., & Laibson, D. (2017). *Myopia and discounting* (National Bureau of Economic Research Working Paper 23254).

Gabaix, X., Laibson, D., Moloche, G., & Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review, 96*(4), 1043–1068.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451–482.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science, 7*(2), 217–229.

Gul, F., & Pesendorfer, W. (2008). Defense of mindless economics. In A. Caplin & A. Schotter (Eds.), *The foundations of positive and normative economics: A handbook* (pp. 193–209). Oxford University Press.

Hadfield-Menell, D., Russell, S. J., Abbeel, P., & Dragan, A. (2016). Cooperative inverse reinforcement learning. *Advances in Neural Information Processing Systems, 29*, 3900–3917.

Handel, B., & Schwartzstein, J. (2018). Frictions or mental gaps: What's behind the information we (don't) use and when do we care? *Journal of Economic Perspectives, 32*(1), 155–178.

Hanna, R., Mullainathan, S., & Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics, 129*(3), 1311–1353.

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory, 104*, 16–47.

Jovanovic, B., & Nyarko, Y. (1995). A Bayesian learning model fitted to a variety of empirical learning curves. *Brookings Papers on Economic Activity. Microeconomics,*, 247–305.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263–292.

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review, 101*(6), 2590–2615.

Kaur, S., Mullainathan, S., Oh, S., & Schilbach, F. (2021). *Do financial concerns make workers less productive?* (National Bureau of Economic Research. Working Paper 28338).

Khaw, M. W., Li, Z., & Woodford, M. (2017). *Risk aversion as a perceptual bias* (National Bureau of Economic Research. Working Paper 23294).

KőSzegi, B., & Rabin, M. (2008). Revealed mistakes and revealed preferences. In A. Caplin & A. Schotter (Eds.), *The foundations of positive and normative economics: A handbook* (pp. 193–209). Oxford University Press.

Kőszegi, B., & Szeidl, A. (2013). A model of focusing in economic choice. *The Quarterly Journal of Economics, 128*(1), 53–104.

Kreiner, C. T., Lassen, D. D., & Leth-Petersen, S. (2014). Measuring the accuracy of survey responses using administrative register data: Evidence from Denmark. *Improving the measurement of consumer expenditures* (pp. 289–307). University of Chicago Press.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics, 112*(2), 443–478.

Lichtenstein, S., & Fischhoff, B. (1980). Training for calibration. *Organizational Behavior and Human Performance, 26*(2), 149–171.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*(1), 46.

Luce, R. D. (1958). A probabilistic theory of utility. *Econometrica, 26*, 193–224.
Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature, 52*(1), 5–44.
Maćkowiak, B., Matějka, F., & Wiederholt, M. (2021). Rational inattention: A review. *Journal of Economic Literature, 61*(1), 226–273.
Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *Quarterly Journal of Economics, 116*(4), 1149–1187.
Malmendier, U., & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics, 126*(1), 373–416.
Malmendier, U., & Nagel, S. (2016). Learning from inflation experiences. *The Quarterly Journal of Economics, 131*(1), 53–87.
Malmendier, U., & Wachter, J. A. (2021). *Memory of past experiences and economic decisions*. Available at SSRN 4013583.
Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science, 341*(6149), 976–980.
Manski, C. F. (2004). Measuring expectations. *Econometrica, 72*(5), 1329–1376.
Manzini, P., & Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review, 97*(5), 1824–1839.
Manzini, P., & Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica, 89*(3), 1153–1176.
Mazzonna, F., & Peracchi, F. (2020). Are older people aware of their cognitive decline? Misperception and financial decision making. *Journal of Political Economy, 132*(6), 1793–1830.
McDonald, A. M., & Cranor, L. F. (2008). The cost of reading privacy policies. *Isjlp, 4*, 543.
Moore, D. A., Swift, S. A., Minster, A., Mellers, B., Ungar, L., Tetlock, P., Yang, H. H., & Tenney, E. R. (2017). Confidence calibration in a multiyear geopolitical forecasting competition. *Management Science, 63*(11), 3552–3565.
Ng, A. Y., & Russell, S. (2000, June). Algorithms for inverse reinforcement learning. *ICML 1*(2), 2.
Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187–192.
Nyarko, Y., & Schotter, A. (2002). An experimental study of belief learning using elicited beliefs. *Econometrica*, *70*(3), 971–1005.
Oprea, R. (2020). What makes a rule complex? *American Economic Review, 110*(12), 3913–3951.
Poterba, J., Venti, S., & Wise, D. (2011). The composition and drawdown of wealth in retirement. *Journal of Economic Perspectives, 25*(4), 95–118.
Puri, I. (2018). *Preference for simplicity*. Available at SSRN 3253494.

Rabin, M., & Thaler, R. H. (2001). Anomalies: Risk aversion. *Journal of Economic Perspectives, 15*(1), 219–232.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*, 873–922.

Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review, 101*(2), 900–926.

Rostow, W. W. (1990). *The stages of economic growth: A non-communist manifesto*. Cambridge University Press.

Rubinstein, A. (1988). Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?) *Journal of Economic Theory, 46*(1), 145–153.

Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., & Aroyo, L. M. (2021, May). "Everyone wants to do the model work, not the data work": Data cascades in high-stakes AI. In *proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–15).

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies, 4*(2), 155–161.

Samuelson, P. A. (1938). A note on the pure theory of consumer's choice. *Economica, 5*, 61–71.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association, 66*(336), 783–801.

Savage, L. J. (1972). *The foundations of statistics*. Dover Publications.

Schotter, A., & Trevino, I. (2014). Belief elicitation in the laboratory. *Annual Review of Economics, 6*(1), 103–128.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593–1599.

Shavell, S. (1995). The appeals process as a means of error correction. *The Journal of Legal Studies, 24*(2), 379–426.

Shiller, R. J. (2020). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

Silver, I., Mellers, B. A., & Tetlock, P. E. (2021). Wise teamwork: Collective confidence calibration predicts the effectiveness of group discussion. *Journal of Experimental Social Psychology, 96*, 104157.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics, 50*, 665–690.

Stephenson, M. C. (2010). Information acquisition and institutional design. *Harvard Law Review, 124*, 1422.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy, 69*(3), 213–225.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies, 23*, 165–180.

Tejeda, H., Kumar, A., Smyth, P., & Steyvers, M. (2022, September). AI-assisted decision-making: A cognitive modeling approach to infer latent reliance strategies. Computational Brain and Behavior. *Carnegie-Rochester conference series on public policy* (Vol. 33, pp. 181–214). North-Holland.

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

Topel, R. (1990). Specific capital and unemployment: Measuring the costs and consequences of job loss. In *Carnegie-Rochester conference series on public policy* (Vol. 33, pp. 181–214). North-Holland.

Toussaert, S. (2018). Eliciting temptation and self-control through menu choices: A lab experiment. *Econometrica, 86*(3), 859–889.

Trueblood, J. S., Li, Y., & DeLosh, E. L. (2021). Human-machine integration in medical decision making. *Current Directions in Psychological Science, 30*(1), 37–43.

Weber, E. H. (1834). *On the tactile senses* (H. E. Ross & D. J. Murray, Trans. and Eds.). Experimental Psychology Series.

Weidmann, B., & Deming, D. J. (2021). Team players: How social skills improve team performance. *Econometrica, 89*(6), 2637–2657.

Welch, F. (1970). Education in production. *Journal of Political Economy, 78*(1), 35–59.

Wiswall, M., & Zafar, B. (2021). Human capital investments and expectations about career and family. *Journal of Political Economy, 129*(5), 1361–1424.

Woodford, M. (2012). Prospect theory as efficient perceptual distortion. *American Economic Review, 102*, 41–46.

Woodford, M. (2020). Modeling imprecision in perception, valuation, and choice. *Annual Review of Economics, 12*, 579–601.

Author Index

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9

Subject Index

A. Caplin, *An Introduction to Cognitive Economics*,
https://doi.org/10.1007/978-3-031-73042-9