

Eliciting probabilistic expectations: Collaborations between psychologists and economists

Wändi Bruine de Bruin^{a,b,1} and Baruch Fischhoff^{b,c}

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved February 6, 2017 (received for review September 18, 2016)

We describe two collaborations in which psychologists and economists provided essential support on foundational projects in major research programs. One project involved eliciting adolescents' expectations regarding significant future life events affecting their psychological and economic development. The second project involved eliciting consumers' expectations regarding inflation, a potentially vital input to their investment, saving, and purchasing decisions. In each project, we sought questions with the precision needed for economic modeling and the simplicity needed for lay respondents. We identify four conditions that, we believe, promoted our ability to sustain these transdisciplinary collaborations and coproduce the research: (i) having a shared research goal, which neither discipline could achieve on its own; (ii) finding common ground in shared methodology, which met each discipline's essential evidentiary conditions, but without insisting on its culturally acquired tastes; (iii) sharing the effort throughout, with common language and sense of ownership; and (iv) gaining mutual benefit from both the research process and its products.

interdisciplinary research | expectations | subjective probabilities | adolescents | inflation

One day in 1995, one of us (B.F.) was standing in line at the National Academy of Sciences Refectory, along with Charles Manski, an economist then at the University of Wisconsin (now at Northwestern University). Manski was complaining about restrictions imposed on the survey questions that he was designing for the 1997 National Longitudinal Study of Youth. To understand adolescents' willingness to invest in their human capital, he wanted to ask them to assess their percent chance of experiencing significant life events that could affect their earning power, such as getting a job as a young adult with or without a high school diploma. However, members of the survey design team maintained that young people could not answer probability questions. They therefore insisted on "easier" response scales, asking adolescents to provide verbal quantifiers, such as "unlikely" and "very likely," instead of percentages. Manski worried that these verbal quantifiers were ambiguous and would elicit responses that could not be compared with actual statistics, making it impossible to tell how well adolescents understood their circumstances. Indeed, Manski saw the constraint as triggering a vicious circle. If he had to ask needlessly vague questions, then the

predictive value of his models would be undermined, making it harder to keep his questions on the survey in the competition for space with other scientists.

Overview

The ensuing conversation was the start of a collaboration between economists and psychologists coproducing research that neither discipline could have done on its own. After summarizing the scientific foundations for eliciting lay expectations, we describe two major collaborative projects, one focused on adolescents' expectations for significant life events and the other on consumers' expectations for inflation. We conclude by describing the conditions that we believe have made this transdisciplinary research possible.

Eliciting Expectations

Existing psychological research confirmed some of Manski's concerns. Psychologists had found that the same verbal quantifier can mean different numbers to different people in a given context and different things to the same person in different contexts (1–4). Thus, a "likely" snowfall in December might imply a higher numerical probability than a "likely"

^aCentre for Decision Research, University of Leeds, Leeds LS2 9JT, United Kingdom; ^bDepartment of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213; and ^cInstitute for Politics and Strategy, Carnegie Mellon University, Pittsburgh, PA 15213

Author contributions: W.B.d.B. and B.F. designed research; W.B.d.B. performed research; W.B.d.B. and B.F. analyzed data; and W.B.d.B. and B.F. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. Email: w.bruinedebruin@leeds.ac.uk.

Table 1. NLSY97 expectation questions, and comparisons with actual outcomes

What is the percent chance that you will	<i>n</i>	Correlation (γ) with observed future outcome	Mean probability (1997), %	Observed future outcome rate, %
Be a student in a regular school a year from now?	3,160	0.64***	92.5	79.6
Have received a high school diploma by the time you turn 20?	3,077	0.60***	94.5	92.0
If you are in school a year from now, [...] be working for pay more than 20 h a week?	2,492	0.29***	57.7	27.2
If you are not in school a year from now, [...] be working for pay more than 20 h a week?	610	0.31***	80.5	43.9
Become pregnant within 1 y from now? (female)	844	0.37***	8.9	20.1
Get someone pregnant within the next year? (male)	1,553	0.35***	9.4	7.9
Become the parent of a baby sometime between now and when you turn 20? (female)	1,368	0.38***	16.0	25.7
Become the parent of a baby sometime between now and when you turn 20? (male)	1,356	0.27***	19.1	13.4
Be arrested, whether rightly or wrongly, at least once in the next year?	3,141	0.41***	10.3	8.2
Serve time in jail or prison between now and when you turn 20?	3,300	0.39***	5.4	2.8
Die from any cause (crime, illness, accident, and so on) in the next year?	3,165	n.s.	18.7	0.1
Die from any cause (crime, illness, accident, and so on) between now and when you turn 20?	3,169	n.s.	20.3	0.5

n.s., not significant. Reprinted from ref. 25, with permission from Elsevier.
*** $P < 0.001$.

snowfall in October (5). Some studies have found that responses to verbal scales predict behaviors and outcomes as well as responses to numerical probability scales (6–8). However, as Manski had observed, verbal scales do not allow responses (which are often provided on 1–7 Likert scales) to be compared with observed probabilities (which vary between 0–100%). As a result, verbal scales could not reveal whether adolescents over- or underestimated their chances of completing school, finding employment, being arrested, or experiencing other significant life events.

In addition to confirming Manski's concerns about verbal quantifiers, studies also had confirmed his hopes about numerical ones. Economists had been asking subjective probability questions on national surveys since the early 1990s, with the University of Michigan's Health and Retirement Study (HRS) being especially influential because of its extensive set of expectation questions, large national sample, and longitudinal design (9). Among its questions, the HRS asked adult respondents to judge the probability that they would live until the age of 75 or 85 y. On average, respondents' probability judgments were close to statistical life table estimates for respondents' age groups and were correlated with their self-reports of known risk factors, such as smoking, drinking, and health conditions (10). Moreover, because the HRS tracked these adults over time, economists were able to find that individuals' judgments of their probability of survival predicted their actual chances of being alive at those target ages (11). Economists also found that HRS respondents who gave lower probabilities of survival were more likely to retire and claim Social Security benefits early (12, 13). In a Dutch household survey that incorporated HRS expectations questions, judged probabilities of stock market gains predicted respondents' stock market participation (9, 14).

The few studies conducted by psychologists examining adolescents' use of numerical probabilities had found similarly

orderly judgments that revealed, among other things, a heightened perception of risk among adolescents as compared with adults (15, 16). One longitudinal study had found that adolescents updated their seemingly high risk judgments after initiating risk behaviors (e.g., drinking, driving), revealing their increased awareness of the relatively lower actual risks that those behaviors entailed (17). Although adolescents' numerical probability judgments sometimes reflected overestimation, the problems seemed to lie with adolescents' underlying beliefs (reflecting, for example, how they interpreted adults' warnings) and not with their ability to express or update those beliefs (18).

Adolescents' Expectations of Significant Life Events

Seeing Manski's plight and an opportunity to extend our own research on adolescent decision making (19, 20), we offered to collaborate. We started by drafting probability questions that captured, in everyday language, the concepts critical to economic models of human capital formation. One of us (W.B.d.B.) then conducted cognitive interviews with adolescents, asking them to think out loud while answering these questions in their own words. The questions were revised, and the process was repeated until our wording was simple enough to be understood by adolescents while remaining precise enough to be useful to economists. The associated instructions, used in the 1997 National Longitudinal Study of Youth (NLSY97), appear in the box below. A subset of the probability questions and responses appear in Table 1. The instructions were longer than we would have preferred, to include details that mattered to Manski and his colleagues. For example, the instructions included explicit encouragement to use nonround numbers, even though respondents rarely did so, except for values close to 0% or 100% (21, 22). It was easier to compromise on the length of the instructions, because the in-person administration of NLSY97 would reduce respondents' cognitive load.

Instructions for the Expectation Module of NLSY97

Each of the next set of questions will ask you for your best guess at the chance that something will happen in the future. You can think of the PERCENT CHANCE that some event will occur as the number of CHANCES OUT OF 100 that the event will take place.

If you think that something is impossible, consider it as having a 0% chance. If you think the event is possible but unlikely, you might say there is a 3% chance or a 15% chance. If you think the chance is pretty even, you can say there is a 46% chance or perhaps a 52% chance. If you think the event is likely, but not certain, you might say there is a 78% chance or a 94% chance. If you think it is certain to happen, give it a 100% chance. Just to make sure that you are comfortable with the scale, I'd like you to do a few practice questions, and explain your answer to me.

What do you think is the percent chance that you will get the flu sometime in the next year?

What do you think is the percent chance that you will eat pizza sometime in the next year?

Source: <https://www.nlsinfo.org/sites/nlsinfo.org/files/attachments/121128/nlsy97r1yexp.html>.

Our collaboration offered a unique opportunity to collect longitudinal survey data for a large representative sample of 15- and 16-y-olds ($n \sim 3,500$) in the United States. The economists on the team offered expertise in the variables critical to human capital formation (e.g., expected return on investment in education) and in the precise terms needed for economic models (e.g., clearly specifying the event and its time of occurrence). The psychologists on the team offered expertise in behavioral decision research and adolescent development [e.g., suggesting heuristics that can both support judgments and lead them astray (23, 24)] and in formulating questions that respondents could understand (e.g., using accessible wording). NLSY97 offered us all a remarkable set of measures of participants' behavior and psychological characteristics, tracked over time, allowing analyses of how respondents' expectations related to their self-reported beliefs and behaviors, for hypotheses posed by both economics and psychology (25–27). Those other measures were created independently by experts in those respective fields. They also reflected different theoretical frames of reference and used diverse response modes, thereby reducing the risk of shared method variance when evaluating the expectation questions (28). In analyses of responses from the initial NLSY97 data, we found sensible relationships between expectations and concurrently reported beliefs and behaviors (thereby reflecting construct validity; ref. 29): For example, adolescents who reported currently being enrolled in school gave higher probabilities for getting a high school diploma by age 20 y (26). In analyses of responses to subsequent waves, we found that reported expectations also predicted future outcomes. For example, adolescents who, in 1997, gave higher probabilities for experiencing various events in the next year or by age 20 y (e.g., "become pregnant," "be arrested, whether rightly or wrongly") were more likely to report having experienced those events by those dates (25). As seen in Table 1, correlations (γ) between expectations and outcomes ranged from 0.27 to 0.64.

For some events (e.g., receiving a high school diploma by age 20 y, being arrested in the next year), teens' mean probability judgment was close to the observed rate, indicating that, on average, they had accurate expectations. For other events, however, expectations were less accurate. For example, adolescents overestimated the probability of being a student a year later and of having a job if not in school. For the psychologists on the team, these results helped to elaborate the processes involved in theories regarding optimistic bias (30), overconfidence (31), and the illusion of control (32). For example, teens' judgment of their probability of being in school a year hence was a poorer predictor than was their current enrollment status, suggesting that some teens underestimate the importance of staying in school for being able to complete their education (33). Nonetheless, their probability judgments predicted future enrollment better than did behavioral measures (e.g., self-reported absenteeism and arrests, or estimates of the percentage of their peers who planned to go to college), suggesting that their expectations summarized knowledge about their lives that was not captured by the behavior self-reports. For the economists on the team, psychological theories provided a framework for interpreting suboptimal judgments and decisions at a time when alternatives to rational actor models were less common in their discipline than they are today.

The NLSY97 responses did have one notable anomaly: adolescents' seemingly pessimistic responses to questions asking for the probability of dying in the next year and by age 20 y (25–27). As seen in the bottom two rows of Table 1, teens greatly overestimated their probability of dying: Although their statistical probability of dying was vanishingly small ($<0.1\%$ annually), their mean judgments were 18% for dying in the next year and 20% for dying by age 20 y. Moreover, responses were similar for the two mortality questions, even though one involved a span of 1 y and the other a span of about 4 y (NLSY97 respondents were, on average, 16 y old at the time of the survey). When respondents gave different probabilities for the two questions, the value was larger for the longer period in only two-thirds of the cases. People may overestimate low probability risks when relying on imperfect knowledge, which makes judgments regressive (34). However, these findings also echoed other evidence that adolescents perceive themselves as especially vulnerable, contradicting the common notion that they have a unique sense of invulnerability (15, 16). Using responses to other NLSY97 modules, we found that teens who gave higher probabilities of dying also reported greater concern about their health and safety and more risk behaviors (27). Responses to later waves of NLSY97 revealed that teens who gave higher probabilities of dying also reported more risk behaviors in young adulthood (35). A similar correlation was observed on the National Longitudinal Study of Adolescent Health, although with a scale using verbal quantifiers that did not allow responses to be compared to statistical estimates (36, 37).

Cognitive interviews, in which we asked participants to think aloud while answering expectation questions, added to our emerging recognition of another psychological aspect of how people think about uncertain events: Sometimes, participants said "50–50," seemingly to express not knowing what to say rather than a numerical probability (38, 39). For example, in one interview, an adolescent responded to a question about the probability of getting HIV from having sex by saying, "It could; it could not. So, I guess it's up in the air ... a 50–50 chance" (38). Another said, "I don't know, 50–50, cause there's always that possibility that something like the condom would break or something" (38).

A review of existing studies on probabilistic expectations revealed a “50 blip” in many response distributions (40, 41). If these 50% responses are intended to mean “I don’t know,” then treating them as numerical values will distort summary statistics. Indeed, there was a substantial 50 blip in responses to the NLSY97 mortality questions, resulting in means (of about 20%) that were much higher than the corresponding medians (of about 10%). In experimental studies pursuing the topic, we found that 50% was more common with open-ended response modes than with visual linear scales (38) and with questions about the probability of a person experiencing an event than with questions about the percent of people experiencing it (39). Working with economist Katherine Carman on the Dutch Longitudinal Internet Panel for the Social Sciences, we found that 50% responses are more likely than other probability responses to be explained as “don’t know” (42).

One notable project building on the NLSY97 collaboration, led by economist Adeline Delavande, has shown the usefulness of eliciting adolescents’ expectations about the effectiveness and side effects of birth control methods in interpreting their choices (43). Delavande and colleagues have also extended research on eliciting expectations in diverse populations. For example, in a sample of 3,000 adults recruited in Malawi (44), she found that respondents’ expectations related to HIV/AIDS were generally coherent, in the sense of following probability axioms and being consistent with their self-reports of related risk behaviors. Despite this internal consistency in their beliefs, respondents reported expectations of dying that were much higher than their actual mortality risk (44). In contrast to these projects, in which psychologists and economists found common ground, there is continuing controversy over how well smokers understand the attendant risks, an area in which researchers from the two disciplines continue to disagree about methods (45, 46).

The NLSY97 expectations module has had modest popularity among economists and other researchers (47–50). However, despite these questions’ construct validity, the seeming overestimation of dying young was troubling for researchers who believed in rational actor models or who were familiar with the much more realistic mortality estimates that adult respondents had provided on other surveys (9, 10). The mortality estimates also were invoked by critics who wanted to discredit the probability response mode altogether, ignoring the evidence of its usefulness. Their complaints were forceful enough to get the expectation module dropped from the next few waves of NLSY97 data collection. Eventually, however, the expectation questions from Wave 1 of NLSY97 proved useful enough that the module was reinstated, although without the mortality questions.

Consumers’ Expectations of Inflation

About 10 y later, Charles Manski called us about a very different expectation question: A team of economic researchers, coordinated by Wilbert van der Klaauw and his colleagues at the Federal Reserve Bank of New York, was studying consumers’ expectations for inflation. For economists and central bankers, these expectations play a central role in predicting financial decisions with implications for the overall economy, for example, about savings, durable goods purchases, and wage negotiations (51). Central banks use surveys to elicit consumers’ inflation expectations, which then, along with market-based measures and surveys of professional forecasters, inform monetary policy, (52, 53). For three decades, the primary survey of inflation expectations in the United States had been the

Reuters/Michigan Survey of Consumers (54). Each month, that survey asked a representative sample of respondents, “During the next 12 months, do you think prices in general will go up, down, or stay where they are now?” Those who answered “up” or “down” were then asked, “By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?” The question was repeated for a longer time horizon, eliciting inflation expectations for the next 5–10 y. It was believed that the wording “prices in general” had been chosen by those who drafted the original survey because it seemed easier for respondents than “inflation,” while having the same meaning for economists.

Respondents’ mean reported expectations for “prices in general over the next 12 mo” had tracked statistical estimates of inflation (55–58). However, the distributions of these responses revealed large variation in individuals’ estimates, with roughly half in the 0–5% range and the remainder including much higher values (54, 56, 59, 60). Because the Consumer Price Index in the United States has not been above 5% since the 1990, such inflation expectations may seem excessive (54, 56). Higher estimates were more common among respondents with lower income and educational attainment (56). One possible economic explanation is that the consumption patterns of individuals in these groups expose them to higher inflation rates, thereby increasing their inflation expectations (61). However, economic analyses revealed much less variation in the inflation rates experienced by different demographic sectors (e.g., from 0.2 to 0.4%, in 1995–2005) than in their inflation expectations (62, 63). One possible psychological explanation is that respondents interpret the expectation question differently.

Social, behavioral, and decision scientists have long studied how people interpret questions (e.g., ref. 64). They have found that even seemingly simple terms, such as “rain,” can mean different things to different people (65). In collaboration with the economists convened by the Federal Reserve Bank of New York, we applied these methods to examine how respondents interpret questions about expectations for “prices in general.” Our first step involved 30 open-ended interviews about financial expectations and decisions, which asked participants to think out loud while answering the standard question about “prices in general.” We found that many did not interpret “prices in general” as referring to overall inflation, as reflected in economic estimates (e.g., the Consumer Price Index) and as had been assumed when survey responses were used in economic models. Rather, some people reported thinking about prices of specific goods, such as gas or food, often focusing on ones that had increased at rates greatly exceeding overall inflation. However, when asked about their expectations for “inflation,” interviewees were more likely to think about overall economic patterns. Contrary to the intuition that the term “inflation” might be unfamiliar, all participants said that they had heard of it and gave reasonable definitions (see also refs. 66 and 67).

The next step of the project examined the generalizability of the interview findings, using RAND’s American Life Panel, a nationally representative sample that regularly completes online surveys. Our first study asked participants how well the question about “prices in general” matched alternative interpretations taken from the interviews. We found that respondents who interpreted “prices in general” as referring to their personal price experiences gave higher expectations for changes in “prices in general” over the next year (59). Our second study found that the personal price experiences most likely to come to

mind were ones involving extreme price changes (68), in line with psychological research showing systematic judgmental biases arising from the salience of extreme experiences (69).

A third study randomly assigned participants to answer questions about their expectations for “prices in general,” “prices you pay,” or “inflation.” Respondents reported expecting greater increases when asked one of the two questions about prices. That effect was mediated by how much they reported thinking about their personal price experiences, which, as mentioned, tended to be relatively high (70). The question about inflation elicited estimates that were closer to economic estimates (e.g., the Consumer Price Index) and more similar across respondents, in terms of both numerical estimates and interpretations. Participants rated the inflation question as more difficult than the ones about prices but were no more likely to skip it (70).*

Thus, our findings suggested that the seemingly simpler wording (i.e., “prices” instead of “inflation”) had needlessly added ambiguity to the expectation question. A follow-up study affirmed the distinction that people make between inflation and prices. Participants who received information about the overall inflation rate were more likely to update their estimates for “the rate of inflation” than were participants who received information about food prices; the opposite was true for participants asked for their expectations about “the prices you pay” (71). Recognizing that the concepts are distinct for respondents allows researchers to pick the wording most appropriate to their theoretical or policy interests. In this case, “inflation” may be better wording for long-term investment decisions, and “prices” may be better wording for nondurable consumer purchases (70).

Another finding from our initial interviews was that respondents often volunteered ranges to express uncertainty about their expectations (e.g., “1–2%”). Although consumer surveys generally elicit point estimates, our economist colleagues were interested in measuring the accompanying uncertainty, because it might help them to understand individual differences in inflation expectations and related decisions. Therefore, we developed a response mode that elicited judgments of the probability of prices falling in specified ranges (e.g., 12% or more, 8–12%, 4–8%, 2–4%, 0–2%, –2 to 0%). Across several surveys, this assessment method appeared meaningful: Almost no participants (1%) skipped the question; at least 89% gave a nonzero probability response to more than one specified range (72). Moreover, respondents who expressed greater uncertainty in this response mode also did so when uncertainty was elicited in other ways: They indicated greater uncertainty on 1–7 rating scales, gave wider ranges around their point estimates, and had point estimates that varied more over time (72).

To test the construct validity of the elicitation procedure, we examined whether these inflation expectations were related to financial decisions (73), as predicted by economic models (74, 75). To avoid inducing consistency artificially, our study asked participants both to make financial decisions and to report their inflation expectations at separate points in the survey (73). The

financial decisions involved choices between an inflation-linked investment and a fixed monetary payment. For each participant, we varied the fixed payment until we found a switching point at which it became more attractive than the inflation-linked investment. As an incentive, we promised participants that two of them would be randomly selected to receive their chosen option. We found that participants’ switching points were monotonically related to their inflation expectations, supporting both the validity of the elicitation procedure and the presumed dependence of expectations and investments (73).

In subsequent research building on these findings, economists have examined how the salience of specific prices can affect overall inflation expectations (76), and psychologists have examined the effects of priming memories of such prices on overall inflation expectations (77). Moreover, our collaboration contributed to the US Federal Reserve launching a new monthly national consumer survey, the Survey of Consumer Expectations, incorporating these jointly developed questions and elicitation methods. It is already being used to inform monetary policy decisions by policy makers in the United States and has inspired a similar survey by the Bank of Canada. The findings of both surveys are publicly available to researchers and practitioners. A full description of our research on inflation expectations is available in a recent review (78).

Conditions for Transdisciplinary Research

The impetus for both research programs was economists’ need for measuring judgments of specific concepts, expressed precisely enough to be compared with actual statistics and used in predictive models. In the first example, the judgments were expectations for significant life events related to human capital formation. In the second example, the judgments were expectations for inflation. In both cases, we followed the same research strategy. First, we learned what each concept meant for economists and how they intended to use it in their models. Second, we conducted interviews, eliciting participants’ mental models of decisions in which the concept might play a role. Third, we designed survey questions, through iterative testing in think-aloud interviews, followed by validation in survey-based experiments.

One concrete measure of the success of these collaborations is the commitment of major longitudinal surveys to using these jointly developed expectation questions. A less concrete measure is the scientific value of the results, which can be judged by the policies and publications using them. In our own research, the collaborations have provided access to large longitudinal national surveys in which responses to our questions could be related to those on modules created by experts in diverse other fields. The NSLY97 collaboration found that adolescents had perhaps surprisingly coherent expectations of future life events, along with theoretically and practically informative biases (e.g., regarding expectations for employment and mortality). The Federal Reserve collaboration found that consumers had perhaps surprising ability to track different parts of the economy and recognize their uncertainty about its future.

In both research programs, the theoretical focus was a technical concept that was thought, by some researchers, to be too complex for lay respondents. In the first example, that concept was probability; in the second, it was inflation. In both cases, it emerged that more useful responses were elicited with the technical term rather than with an alternative intended to make the task easier. In the first example, numerical probabilities revealed

*In a more recent study with a Dutch panel, we studied the effect of (web vs. face-to-face) administration mode on responses to the “inflation” question and the “prices in general” question (89). We replicated the finding that non-response rates were similar for both questions in the web survey but were much higher for the “inflation” question in face-to-face mode. These increased non-response rates might reflect interviewees’ greater hesitance when answering questions that they perceive as more difficult in front of an interviewer and interviewers’ greater tendency to permit skipping questions they perceive as more difficult.

construct and predictive validity, with the precision needed to inform economic models and evaluate the accuracy of lay beliefs (compared with statistical estimates). In the second example, "inflation" proved to be a meaningful term, evoking consumers' thinking about the economy as a whole, whereas ostensibly simpler terms invoking prices directed respondents to salient, unrepresentative changes.

Both research programs involved transdisciplinary efforts. We believe that the conditions allowing that collaboration included:

- i) Shared research goal. Both projects were motivated by a common goal: designing survey questions that would improve understanding of individuals' expectations relevant to their personal and financial welfare. All team members agreed in advance that those questions should have the specificity required by economic models and the validity required by psychological methods, recognizing the essential contribution of both disciplines.
- ii) Shared methodology. Where economic and psychological approaches diverged, we sought common ground. We set aside differences in research tastes without theoretical or practical implications. We subjected differences that did matter to empirical test (e.g., think-aloud interviews to examine respondents' question interpretations, experimental comparisons of alternative approaches, correlations evaluating construct or predictive validity).
- iii) Shared effort. On both projects, the teams collaborated throughout the process, taking care that one another's concerns were properly understood and adequately addressed, and that agreement was reached on how to proceed. That continuous consultation and coproduction allowed the teams to develop shared language and ownership. In fact, our efforts were shared to the extent that, in writing this paper, we often found it difficult to disentangle precisely who had made which specific contributions.
- iv) Shared benefits. Both projects afforded benefits that participants could not have found working solely within their discipline. For the economists, those benefits included psychologists' insight into question construction and validation, along with evidence supporting their desire for better survey methods. For the psychologists, those benefits included access to unique data sets and to economists' formalization of how people might form expectations and use them in

decisions regarding human capital formation and personal finance.[†]

Achieving these conditions involved a shared conviction that the research required more than just reading relevant publications in one another's literature. Rather, it required continuing consultation with colleagues in the other discipline who knew the context within which those publications were written. It also involved acknowledging when the literature contains competing views and evidence, open issues, and traditions based on taste rather than evidence.

We have found these conditions met in other teams of psychologists and economists willing to make similar personal commitments to understanding personal and social decisions (42, 79). We also have been fortunate enough to find them in collaborations with medical experts, on understanding and informing patients' health decisions (80); with engineers, in analyzing risks (81, 82) and promoting sustainable consumer decisions (83, 84); with environmental scientists, in understanding public response to extreme weather and climate change (85, 86); and with anthropologists and political scientists in studying security issues (87, 88). We believe that such collaborations are essential for addressing applied problems and are healthy for the sciences involved, by testing their theories in complex environments and revealing phenomena that stimulate new research. We have also found them personally and intellectually gratifying.

Acknowledgments

We thank our collaborators for inspiring us, especially Charles Manski and Wilbert van der Klaauw, who also provided feedback on an earlier version of this paper. Financial support for preparing this paper was provided in part by grants from the Swedish Foundation for the Humanities and the Social Sciences (Riksbankens Jubileumsfond) Program on Science and Proven Experience and by the European Union Seventh Framework Program PEOPLE-2013-CIG-618522). The NLSY97 was sponsored by the Bureau of Labor Statistics at the US Department of Labor.

[†]We believe that these collaborations also met the criteria for creative research environments identified by Nils-Eric Sahlin (90): (i) generosity, in sharing knowledge and experience; (ii) community; (iii) scientific qualifications; (iv) intellectual diversity; (v) trust and tolerance for ideas breaking with traditions; (vi) equality, among participants of varying seniority; (vii) curiosity; (viii) freedom of spirit; and (ix) intimacy of scale, allowing individuals to stay in touch.

- 1 Beyth-Marom R (1982) How probable is probable? A numerical translation of verbal probability expressions. *J Forecast* 1:257–269.
- 2 Brun W, Teigen KH (1988) Verbal probabilities: Ambiguous, context-dependent, or both? *Organ Behav Hum Decis Process* 41:390–404.
- 3 Budescu DV, Wallsten TS (1985) Consistency in interpreting probabilistic phrases. *Organ Behav Hum Decis Process* 36:391–405.
- 4 Lichtenstein S, Newman JR (1967) Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychon Sci* 9:563–564.
- 5 Teigen HV, Brun W (2003) Verbal expressions of uncertainty and probability. *Thinking. Psychological Perspectives on Reasoning, Judgement, and Decision Making*, eds Hardman D, Macchi L (Wiley, New York), pp 125–146.
- 6 Weinstein ND, Diefenbach MA (1997) Percentage and verbal category measures of risk likelihood. *Health Educ Res* 12:139–141.
- 7 Windschitl PD, Wells GL (1996) Measuring psychological uncertainty: Verbal versus numeric methods. *J Exp Psychol Appl* 2:343–364.
- 8 Woloshin S, Schwartz LM, Byram S, Fischhoff B, Welch HG (2000) A new scale for assessing perceptions of chance: A validation study. *Med Decis Making* 20(3): 298–307.
- 9 Hurd MD (2009) Subjective probabilities in household surveys. *Annu Rev Econ* 1:543–562.
- 10 Hurd MD, McGarry K (1995) Evaluation of the subjective probabilities of survival in the Health and Retirement Study. *J Hum Resour* 30:s268–s292.
- 11 Hurd MD, McGarry K (2002) The predictive validity of subjective probabilities of survival. *Econ J (Oxf)* 112:966–985.
- 12 Hurd MD, Smith JP, Zissimopoulos JM (2004) The effects of survival on retirement and Social Security claiming. *J Appl Econ* 19:761–775.
- 13 Delavande A, Perry M, Willis RJ (2006) *Probabilistic Thinking and Early Social Security Claiming*. Working Paper 2006-129 (Michigan Retirement Research Center, University of Michigan, Ann Arbor, MI).
- 14 Hurd MD, van Rooij M, Winter J (2011) Stock market expectations of Dutch households. *J Appl Econ (Chichester Engl)* 26(3):416–436.
- 15 Millstein SG, Halpern-Felsher BL (2002) Judgments about risk and perceived invulnerability in adolescents and young adults. *J Res Adolesc* 12:399–422.
- 16 Quadrel MJ, Fischhoff B, Davis W (1993) Adolescent (in)vulnerability. *Am Psychol* 48(2):102–116.
- 17 Gerrard M, Gibbons FX, Benthin AC, Hessling RM (1996) A longitudinal study of the reciprocal nature of risk behaviors and cognitions in adolescents: What you do shapes what you think, and vice versa. *Health Psychol* 15(5):344–354.

- 18 Halpern-Felsher BL, et al. (2001) The role of behavioral experience in judging risks. *Health Psychol* 20(2):120–126.
- 19 Fischhoff B (1996) The real world: What good is it? *Organ Behav Hum Decis Process* 65:232–248.
- 20 Fischhoff B, Downs JS, Bruine de Bruin W (1998) Adolescent vulnerability: A framework for behavioral interventions. *Appl Prev Psychol* 7:77–94.
- 21 Dominitz J, Manski CF (1997) Perceptions of economic insecurity: Evidence from the Survey of Economic Expectations. *Public Opin Q* 61:261–287.
- 22 Manski CF, Molinari F (2010) Rounding probabilistic expectations in surveys. *J Bus Econ Stat* 28(2):219–231.
- 23 Kahneman D (2011) *Thinking Fast and Slow* (Farrar, Strauss & Giroux, New York).
- 24 Tversky A, Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science* 185(4157):1124–1131.
- 25 Bruine de Bruin W, Parker AM, Fischhoff B (2007) Can adolescents predict significant life events? *J Adolesc Health* 41(2):208–210.
- 26 Fischhoff B, et al.; Bruine de Bruin W (2000) Teen expectations for significant life events. *Public Opin Q* 64(2):189–205.
- 27 Fischhoff B, Bruine de Bruin W, Parker AM, Millstein SG, Halpern-Felsher BL (2010) Adolescents' perceived risk of dying. *J Adolesc Health* 46(3):265–269.
- 28 Podsakoff PM, MacKenzie SB, Podsakoff NP (2012) Sources of method bias in social science research and recommendations on how to control it. *Annu Rev Psychol* 63:539–569.
- 29 Cronbach LJ, Meehl PE (1955) Construct validity in psychological tests. *Psychol Bull* 52(4):281–302.
- 30 Weinstein ND (1980) Unrealistic optimism about future life events. *J Pers Soc Psychol* 39:806–820.
- 31 Koriart A, Lichtenstein S, Fischhoff B (1980) Reasons for confidence. *J Exp Psychol Human* 6:107–118.
- 32 Langer EJ (1975) The illusion of control. *J Pers Soc Psychol* 32:311–328.
- 33 Persoskie A (2003) How well can adolescents really judge risk? Simple, self-reported risk factors out-predict teens' self-estimates of personal risk. *Judgm Decis Mak* 8:1–6.
- 34 Moore DA, Healy PJ (2008) The trouble with overconfidence. *Psychol Rev* 115(2):502–517.
- 35 McDade TW, et al. (2011) Adolescents' expectations for the future predict health behaviors in early adulthood. *Soc Sci Med* 73(3):391–398.
- 36 Duke NN, Skay CL, Pettingell SL, Borowsky IW (2009) Adolescent perception of premature risk for death: Contributions from individual and environmental contextual factors. *Acad Pediatr* 9(4):256–262.
- 37 Duke NN, Borowsky IW, Pettingell SL, Skay CL, McMorris BJ (2011) Adolescent early death perception: Links to behavioral and life outcomes in young adulthood. *J Pediatr Health Care* 25(4):224–234.
- 38 Fischhoff B, Bruine de Bruin W (1999) Fifty-fifty=50%? *J Behav Decis Making* 12:149–163.
- 39 Bruine de Bruin W, Fischhoff B, Millstein SG, Halpern-Felsher BL (2000) Verbal and numerical expressions of probability: "It's a fifty-fifty chance." *Organ Behav Hum Decis Process* 81(1):115–131.
- 40 Dominitz J, Manski CF (1997) Using expectations data to study subjective income expectations. *J Am Stat Assoc* 92:855–867.
- 41 Viscusi WK (1993) *The Risk of Smoking* (Harvard Univ Press, Cambridge, MA).
- 42 Bruine de Bruin W, Carman KG (2012) Measuring risk perceptions: What does the excessive use of 50% mean? *Med Decis Making* 32(2):232–236.
- 43 Delavande A (2008) Pill, patch, or shot? Subjective expectations and birth control choice. *Int Econ Rev* 49:999–1042.
- 44 Delavande A, Kohler HP (2009) Subjective expectations in the context of HIV/AIDS in Malawi. *Demogr Res* 20:817–874.
- 45 Slovic P, ed (2001) *Smoking: Risk, Perception and Policy* (Sage, Thousand Oaks, CA).
- 46 Viscusi WK (2002) *Smoke-Filled Rooms* (Univ of Chicago Press, Chicago).
- 47 Dominitz J, Manski CF, Fischhoff B (2001) Who are youth "at risk"? Expectations evidence in the NLSY97. *Social Awakenings: Adolescents' Behavior as Adulthood Approaches*, ed Michael RT (Russell Sage Foundation, New York), pp 230–257.
- 48 Gassanov MA, Nicholson LM, Koch-Turner A (2008) Expectations to marry among American youth: The effects of unwed fertility, economic activity, and cohabitation. *Youth Soc* 40:265–288.
- 49 Reynolds JR, Pemberton J (2001) Rising college expectations among youth in the United States: A comparison of the 1979 and 1997 NLSY. *J Hum Resour* 36:703–726.
- 50 Sweeten G (2006) Who will graduate? Disruption of high school education by arrest and court involvement. *Justice Q* 23:462–480.
- 51 Katona G (1975) *Psychological Economics* (Elsevier, Oxford, UK).
- 52 Cunningham R, Desroches B, Santor E (2010) Inflation expectations and the conduct of monetary policy: A review of recent evidence and experience. *Bank of Canada Review, Spring*, 13–25.
- 53 Gali J (2008) *Monetary policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework* (Princeton Univ Press, Princeton, NJ).
- 54 Curtin RT (1996) *Procedure to Estimate Price Expectations* (Univ of Michigan Survey Research Center, Ann Arbor, MI).
- 55 Ang A, Bekaert G, Wei M (2007) Do macro variables, asset markets or surveys forecast inflation better? *J Monet Econ* 54:1163–1212.
- 56 Bryan MF, Venkatu G (2001) The demographics of inflation opinion surveys. *Federal Reserve Bank of Cleveland Economic Commentary* 10:1–4.
- 57 Hafer RW, Hein SE (1985) On the accuracy of time-series, interest rate, and survey forecasts of inflation. *J Bus* 58:377–398.
- 58 Thomas LB, Jr (1999) Survey measures of expected US inflation. *J Econ Perspect* 13:125–144.
- 59 Bruine de Bruin W, et al. (2010) Expectations of inflation: The role of financial literacy and demographic variables. *J Consum Aff* 44:381–402.
- 60 Mankiw NG, Reis R, Wolfers J (2004) Disagreement about inflation expectations. *NBER Macroeconomics Annual* 2003(18):209–268.
- 61 McGranahan L, Paulson A (2006) *Constructing the Chicago Fed Income Based Economic Index – Consumer Price Index: Inflation expectations by Demographic Group: 1983-2005*. Federal Reserve Bank of Chicago Working Paper 2005-20 (Federal Reserve Bank of Chicago, Chicago).
- 62 Hobijn B, Mayer K, Stennis C, Topa G (2009) *Household Inflation Experience in the US: A Comprehensive Approach*. Federal Reserve Bank of San Francisco Working Paper 2009-19 (Federal Reserve Bank of San Francisco, San Francisco).
- 63 Kokoski M (2000) Alternative CPI aggregations: Two approaches. *Mon Labor Rev* 123:31–39.
- 64 Turner CF, Martin E (1984) *Surveying Subjective Phenomena* (Russell Sage Foundation, New York).
- 65 Murphy AH, Lichtenstein S, Fischhoff B, Winkler RL (1980) Misinterpretations of precipitation probability forecasts. *Bull Am Meteorol Soc* 61:695–701.
- 66 Leiser D, Drori S (2005) Naïve understanding of inflation. *J Socio-Economics* 34:179–198.
- 67 Svenson O, Nilsson G (1986) Mental economics: Subjective representations of factors related to expected inflation. *J Econ Psychol* 7:327–349.
- 68 Bruine de Bruin W, van der Klaauw W, Topa G (2011) Inflation expectations: The biasing effect of thoughts about specific prices. *J Econ Psychol* 32:834–845.
- 69 Tversky A, Kahneman D (1973) Availability: A heuristic for judging frequency and probability. *Cognit Psychol* 4:207–232.
- 70 Bruine de Bruin W, et al. (2012) The effect of question wording on consumers' reported inflation expectations. *J Econ Psychol* 4:749–757.
- 71 Armantier O, Nelson S, Topa G, van der Klaauw W, Zafar B (2016) The price is right: Updating inflation expectations in a randomized price information experiment. *Rev Econ Stat* 98:503–523.
- 72 Bruine de Bruin W, Manski CF, Topa G, van der Klaauw W (2011) Measuring consumer uncertainty about future inflation. *J Appl Econ* 26:454–478.
- 73 Armantier O, Bruine de Bruin W, Topa G, van der Klaauw W, Zafar B (2015) Inflation expectations and behavior: Do survey respondents act on their beliefs? *Int Econ Rev* 56:505–536.
- 74 Sims C (2009) *Inflation Expectations, Uncertainty and Monetary Policy*. Working Paper 275 of the Bank for International Settlements (Bank for International Settlements, Basel) Souleles N. 2.

- 75 Woodford M (2005) Central bank communication and policy effectiveness. *Proceedings of the Federal Reserve Bank of Kansas City* 2005(Aug):399–474.
- 76 Cavallo A, Cruces G, Perez-Truglia R, Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *Am Econ J Macroecon*, in press.
- 77 Del Missier F, Ranyard R, Bonini N (2016) Perceived inflation: The role of product accessibility and attitudes towards inflation. *J Econ Psychol* 56:97–106.
- 78 Armantier O, et al. (2013) Measuring inflation expectations. *Annu Rev Econ* 5:273–301.
- 79 Parker AM, Bruine de Bruin W, Yoong J, Willis R (2012) Inappropriate confidence and retirement planning: Four studies with a national sample. *J Behav Decis Making* 25(4):382–389.
- 80 Downs JS, et al. (2004) Interactive video behavioral intervention to reduce adolescent females' STD risk: A randomized controlled trial. *Soc Sci Med* 59(8): 1561–1572.
- 81 Fischhoff B (2015) The realities of risk-cost-benefit analysis. *Science* 350(6260):aaa6516.
- 82 Florig HK, Fischhoff B (2007) Individuals' decisions affecting radiation exposure after a nuclear explosion. *Health Phys* 92(5):475–483.
- 83 Mayer LA, Bruine de Bruin W, Morgan MG (2014) Informed public choices for low-carbon electricity portfolios using a computer decision tool. *Environ Sci Technol* 48(7):3640–3648.
- 84 Min J, Azevedo IL, Michalek J, Bruine de Bruin W (2014) Labeling energy costs on light bulbs lowers implicit discount rates. *Ecol Econ* 97:42–50.
- 85 Bruine de Bruin W, et al. (2016) Promoting protection against a threat that evokes positive affect: The case of heat waves in the United Kingdom. *J Exp Psychol Appl* 22(3):261–271.
- 86 Taylor A, Bruine de Bruin W, Dessai S (2014) Climate change beliefs and perceptions of weather-related changes in the United Kingdom. *Risk Anal* 34(11): 1995–2004.
- 87 Atran S, Axelrod R, Davis R, Fischhoff B (2017) Challenges in researching terrorism from the field. *Science* 355(6323):352–354.
- 88 Fischhoff B, Atran S, Fischhoff N (2007) Counting casualties: A framework for respectful, useful records. *J Risk Uncertain* 34:1–19.
- 89 Bruine de Bruin W, van der Klaauw W, van Rooij M, Teppa F, de Vos K Measuring expectations of inflation: Effects of survey mode, wording, and opportunities to revise. *J Econ Psychol*, in press.
- 90 Sahlin NE (2013) Creating creative research environments. *Trust and Confidence in Scientific Research*, eds Hermerén G, Sahlin K, Sahlin NE (The Royal Swedish Academy of Letters, History and Antiquities, Stockholm), pp 80–87.