This is a repository copy of *Eliciting probabilistic expectations: Collaborations between psychologists and economists*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/112039/

Version: Accepted Version

**Article:**
Bruine de Bruin, W orcid.org/0000-0002-1601-789X and Fischhoff, B (2017) Eliciting probabilistic expectations: Collaborations between psychologists and economists. Proceedings of the National Academy of Sciences, 114 (13). pp. 3297-3304. ISSN 1091-6490

https://doi.org/10.1073/pnas.1615461114

This is an author produced version of a paper published in Proceedings of the National Academy of Sciences. Uploaded in accordance with the publisher's self-archiving policy. In order to comply with the publisher requirements the University does not require the author to sign a non-exclusive licence for this paper.

**Reuse**
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Eliciting Probabilistic Expectations:
Collaborations between Psychologists and Economists

Proceedings of the National Academy of Sciences (PNAS), in press

Wändi Bruine de Bruin$^1,2$
Baruch Fischhoff$^2,3$

$^1$Centre for Decision Research, University of Leeds, UK
$^2$Department of Engineering and Public Policy, Carnegie Mellon University, US
$^3$Institute for Politics and Strategy, Carnegie Mellon University, US
Abstract

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: (1) having a shared research goal, which neither discipline could achieve on its own; (2) finding common ground in shared methodology, which met each discipline’s essential evidentiary conditions, but without insisting on its culturally acquired tastes; (3) sharing the effort throughout, with common language and sense of ownership; and (4) gaining mutual benefit from both the research process and its products.

Keywords: interdisciplinary research, expectations, economics, psychology
Introduction

One day in 1995, one of us (BF) 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). Manski was complaining about restrictions imposed on the survey questions that he was designing for the 1997 National Longitudinal Study of Youth. In order 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, the survey’s directors 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 to 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, which would then make it harder to keep his questions on the survey – in the fight for space with other scientists.

Overview. The ensuing conversation was the start of a collaboration between economists and psychologists co-producing 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
Collaborating on probabilistic expectations 4

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 to the same person in different contexts (1, 2, 3, 4). Thus, a ‘likely’ snowfall in December might imply a higher numerical probability than a ‘likely’ 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, 7, 8). However, as Manski had observed, verbal scales do not allow comparing responses (which are often provided on 1-7 Likert scales) to 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 had also 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 due to its extensive set of expectations questions, large national sample, and longitudinal design (9). Among its questions, HRS asked adult respondents to judge the probability that they would live until the ages 75 or 85. On average, respondents’ probability judgments were close to statistical life table estimates for respondents’ age groups, and were correlated with their self-reports on known risk factors, such as smoking, drinking, and health conditions (10). Moreover, because the HRS tracked these
adults over time, economists were able to show 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 examining adolescents’ use of numerical probabilities, conducted by psychologists, had found similarly orderly judgments, which revealed, among other things, a heightened perception of risk among adolescents, as compared to 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 overestimations, the problems seemed to lie with their 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 (WBB) 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 repeated, until our wording was simple enough to be understood by
adolescents, while remaining precise enough to be useful to economists. The resulting
instructions, used in the 1997 National Longitudinal Study of Youth (NLSY97), appear in Table
1. A subset of the probability questions and responses appear in Table 2. The instructions were
longer than we would have preferred, in order to include details that mattered to Manski and his
colleagues. For example, the instructions included explicit encouragement to use non-round
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.

Our collaboration offered a unique opportunity to collect longitudinal survey data for a
large representative sample of US 15- and 16-year olds (N≈3500). The economists on the team
offered expertise in the variables critical to human capital formation (e.g., expected return on
investment in education), as well as 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]), as well as 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, 26,
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 of
expectations with concurrently reported beliefs and behaviors (or construct validity; 29): For example, adolescents who reported currently being enrolled in school gave higher probabilities for getting a high-school diploma by age 20 (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 were more likely to report having experienced them by those dates (e.g., “become pregnant,” “be arrested, whether rightly or wrongly”) (25). As seen in Table 2, correlations ($\gamma$) between expectations and outcomes ranged from .27 to .64.

For some events, teens’ mean probability judgment was close to the observed rate (e.g., receiving a high school diploma by age 20, being arrested in the next year), indicating that, on average, they had accurate expectations. For other events, though, 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’ judgments of their probability of being in school a year hence was a poorer predictor than was their current enrollment, 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, such as 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, reflected in adolescents’ seemingly pessimistic responses to questions asking for the probability of dying in the next year and by age 20 (25, 26, 27). As seen in the bottom two rows of Table 2, 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. Moreover, responses were similar for the two questions, even though one involved a year and the other about four years (as NLSY97 respondents were, on average, 16 years old at the time). When respondents gave different probabilities for the two questions, the value was larger for the longer period in only two-thirds of the cases. These findings echoed other evidence of adolescents perceiving themselves as especially vulnerable — contradicting the common wisdom 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 (34). (A similar correlation was observed on the National Longitudinal Study of Adolescent Health, although with a scale using verbal quantifiers that did not allow comparing responses to statistical estimates [35, 36]).

Cognitive interviews, in which we asked participants to think aloud while answering expectations 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 (37, 38). 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 fifty-fifty chance” (37). Another said, “I don’t know, fifty-fifty, cause there’s always that possibility that something like the condom would break or something” (37).

A review of existing studies on probabilistic expectations revealed a “50 blip” in many response distributions (39, 40). 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 (37), and with questions about the probability of a person experiencing an event than with questions about the percent of people experiencing it (38). Working with economist Katherine Carman on the Dutch Longitudinal Internet Panel for the Social Sciences (LISS), we found that 50% responses are more likely than other probability responses to be explained as “don’t know” (41).

One notable project building on the NLSY97 collaboration, led by economist Adeline Delavande, has shown the usefulness of eliciting adolescents’ expectations about effectiveness and side effects of birth control methods, for interpreting their choices (42). 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 (43), 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 (43). In contrast to these projects, where psychologists and economists found common ground, there is continuing controversy over how well smokers understand the attendant risks, where researchers from the two disciplines did not (44, 45).

The NLSY97 expectations module has had modest popularity among economists and other researchers (46, 47, 48,49). Yet, 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 were also 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, though, 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 years 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 (50). Central banks use surveys to elicit consumers’ inflation expectations, which then inform monetary policy, along with market-based measures and surveys of professional forecasters (51, 52). For three decades, the primary U.S.
survey of inflation expectations had been the Reuters/Michigan Survey of Consumers (53). Each month, that survey asked a representative sample of U.S. 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 also repeated for a longer time horizon, eliciting inflation expectations for the next 5 to 10 years. The wording “prices in general” had, it was believed, 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 months had tracked statistical estimates of inflation (54, 55, 56, 57). 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 (53, 55, 58, 59). Because the Consumer Price Index in the U.S. has not been above 5% since the 1990, such inflation expectations may seem excessive (53, 55). Higher estimates were more common among respondents with lower income and educational attainment (55). 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 (60). However, economic analyses revealed much less variation in the inflation rates experienced by different demographic sectors (e.g., from .2-.4%, in 1995-2005) than in their inflation expectations (61, 62). 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., 63). They have found that even seemingly simple terms, such as “rain,” can
mean different things to different people (64). 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 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.
When asked about their expectations for “inflation,” though, 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 65, 66).

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 (58). Our second study found
that the personal price experiences most likely to come to mind were ones involving extreme
price changes (67), in line with psychological research showing systematic judgmental biases
due to the salience of extreme experiences (68).
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 personal price experiences – which, as mentioned, tended to be relatively high (69). The “inflation” question 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 (69).²

Thus, our findings suggested that the seemingly simpler wording (i.e. “prices” instead of “inflation”) had needlessly added ambiguity to the expectations question. A follow-up study affirmed the distinction that people make between inflation versus 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” (70). 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’ better for nondurable consumer purchases.

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 surrounding uncertainty, as 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% to 12%, 4% to 8%, 2% to 4%, 0% to 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 (71). 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 a 1-7 rating scales, gave wider ranges around their point estimates, and had point estimates that varied more over time (71).

To test the construct validity of the elicitation procedure, we examined whether these inflation expectations were related to financial decisions (72), as predicted by macro-economic models (73, 74). Our study asked participants both to make financial decisions and to report their inflation expectations – at separate points in the survey, so as not to induce consistency artificially (72). The financial decisions involved choices between an inflation-linked investment and a fixed monetary payment. For each participant, we varied that fixed payment until we found a switching point, where 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 between expectations and investments (72).

In subsequent research building on these findings, economists have examined how the salience of specific prices can affect overall inflation expectations (75), while psychologists have examined the effects of priming memories of such prices on overall inflation expectations (76). Moreover, our collaboration contributed to the U.S. 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 U.S. policy makers, 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 (77).

Conditions for Transdisciplinary Research

Summary. The impetus for both research programs was economists’ need for measuring
judgments of specific concepts, expressed precisely enough to be compared to 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 new 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, where 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 an alternative intended to make the task easier. In the first example, numerical probabilities revealed construct and predictive validity, with the precision needed to inform economic models and evaluate the accuracy of lay beliefs (compared to 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:

1. 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.

2. 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 about respondents’ question interpretations, experimental comparisons of alternative approaches, correlations evaluating construct or predictive validity).

3. 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 co-production 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.

4. Shared benefits. Both projects afforded benefits that participants could not have found working solely within their discipline. For the economists, those benefits included 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 are 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 (41, 78). We have also been fortunate enough to find them in collaborations with medical experts, on understanding and informing patients’ health decisions (79); with engineers, in
analyzing risks (80, 81) and promoting sustainable consumer decisions (82, 83); with environmental scientists, in understanding public response to extreme weather and climate change (84, 85); and with anthropologists and political scientists in studying security issues (86, 87). We believe that such collaborations are essential to addressing applied problems and 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.

Footnotes

1 People may also overestimate low-probability risks, because of relying on imperfect knowledge, which makes judgments regressive (88).

2 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.

3 We believe that these collaborations also met the criteria for creative research environments identified by Nils-Eric Sahlin (90): (1) generosity, in sharing knowledge and experience; (2) community, (3) scientific qualifications; (4) intellectual diversity; (5) trust and tolerance, for
ideas breaking with traditions; (6) equality, among participants of varying seniority; (7) curiosity; (8) freedom of spirit; (9) intimacy of scale, allowing individuals to stay in touch.

Acknowledgments

Financial support for preparing this manuscript 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 (FP7-PEOPLE-2013-CIG-618522). The NLSY97 was sponsored by the Bureau of Labor Statistics at the U.S. Department of Labor. We thank our collaborators for inspiring us, especially Chuck Manski and Wilbert van der Klaauw, who also provided feedback on an earlier version of this manuscript.

References


Table 1 Expectation Instructions for 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 percent chance. If you think the event is possible but unlikely, you might say there is a 3 percent chance or a 15 percent chance. If you think the chance is pretty even, you can say there is a 46 percent chance or perhaps a 52 percent chance. If you think the event is likely, but not certain, you might say there is a 78 percent chance or a 94 percent chance. If you think it is certain to happen, give it a 100 percent 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
Table 2. NLSY97 expectations questions, and comparisons with actual outcomes.

<table>
<thead>
<tr>
<th>What is the percent chance that you will..</th>
<th>N</th>
<th>Correlation (γ) with observed future outcome</th>
<th>Mean probability (1997)</th>
<th>Observed future outcome rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be a student in a regular school a year from now?</td>
<td>3160</td>
<td>.64***</td>
<td>92.5%</td>
<td>79.6%</td>
</tr>
<tr>
<td>Have received a high school diploma by the time you turn 20?</td>
<td>3077</td>
<td>.60***</td>
<td>94.5%</td>
<td>92.0%</td>
</tr>
<tr>
<td>If you are in school a year from now, [...] working for pay more than 20 hours a week?</td>
<td>2492</td>
<td>.29***</td>
<td>57.7%</td>
<td>27.2%</td>
</tr>
<tr>
<td>If you are not in school a year from now, [...] working for pay more than 20 hours a week?</td>
<td>610</td>
<td>.31***</td>
<td>80.5%</td>
<td>43.9%</td>
</tr>
<tr>
<td>Become pregnant within 1 year from now? (female)</td>
<td>844</td>
<td>.37***</td>
<td>8.9%</td>
<td>20.1%</td>
</tr>
<tr>
<td>Get someone pregnant within the next year? (male)</td>
<td>1553</td>
<td>.35***</td>
<td>9.4%</td>
<td>7.9%</td>
</tr>
<tr>
<td>Become the parent of a baby sometime between now and when you turn 20? (female)</td>
<td>1368</td>
<td>.38***</td>
<td>16.0%</td>
<td>25.7%</td>
</tr>
<tr>
<td>Become the parent of a baby sometime between now and when you turn 20? (male)</td>
<td>1356</td>
<td>.27***</td>
<td>19.1%</td>
<td>13.4%</td>
</tr>
<tr>
<td>Be arrested, whether rightly or wrongly, at least once in the next year?</td>
<td>3141</td>
<td>.41***</td>
<td>10.3%</td>
<td>8.2%</td>
</tr>
<tr>
<td>Serve time in jail or prison between now and when you turn 20?</td>
<td>3300</td>
<td>.39***</td>
<td>5.4%</td>
<td>2.8%</td>
</tr>
<tr>
<td>Die from any cause (crime, illness, accident, and so on) in the next year?</td>
<td>3165</td>
<td>n.s.</td>
<td>18.7%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Die from any cause (crime, illness, accident, and so on) between now and when you turn 20?</td>
<td>3169</td>
<td>n.s.</td>
<td>20.3%</td>
<td>0.5%</td>
</tr>
</tbody>
</table>

*** p<.001; n.s.=not significant

Source: Bruine de Bruin et al. (2007)