16

Ethnographic Decision Models with Qualitative Data

A Thoroughly Mixed Method[1]

Gery W. Ryan and H. Russell Bernard

Introduction: Definition

Ethnographic decision models (EDMs) are a multimethod approach that predicts episodic behaviors, like choosing among alternative treatments during illness (Ryan and Martinez 1996) or deciding whether to evacuate during a hurricane (Gladwin et al. 2001). EDMs are based on a deep understanding of the cultural rules and nuances that people use to make decisions. As an approach, EDM consists of an ordered, rigorous set of data collection and analysis techniques. To start, researchers use semi-structured interviews to collect qualitative data on people's decision-making process and then use qualitative methods to identify decision criteria. Next, researchers develop and administer a survey instrument to collect more structured—i.e., quantitative—data on people's actual choices and the circumstances under which they made those choices. These survey data are used to build an ethnographically grounded decision model that anticipates or accounts for the decision outcomes. Constructing the initial model can be done by building from one case to another—a thoroughly qualitative approach—or it can be done using machine learning—a quantitatively intensive approach (Murthy 1998; Zhang 2016). Finally, researchers test the accuracy of the model quantitatively, with a new sample, by estimating the degree to which it predicts the reported actual outcomes at or better than chance and qualitatively by testing its content validity (e.g., to what degree does the model conform to how participants describe their own decision processes). Because EDM inherently relies on both qualitative and quantitative approaches, it is classified in this book as an "Inherently" Mixed Analysis Approach.

When Use of EDMs Is Appropriate in MMR

EDMs are used when a researcher wants to understand individual decision-making processes leading to outcomes of interest. Any recurring decision—to buy or not buy a computer; to use (or demand the use of) a condom during sex; to stay home sick from work or not—can be modeled with the EDM method. As with any binary outcome, yes–no decisions can be modeled and tested with logistic regression and there are computer programs for modeling decisions, but crucially, ethnographic decision models illuminate *how choices are made from the perspective of those making those choices.* To generate and test EDMs, researchers use a multi-staged approach to combine many of the exploratory data collection and preliminary model building techniques used in grounded theory (Corbin and Strauss 2015; Glaser and Strauss 1967) with the confirmatory and validating techniques used in classic content analysis (Krippendorf 2013). Typically, EDMs are built and tested on small, ethnographic samples of interviews with 20–60 people with one sample being used to build the model and another independent sample used to test the model. Even with such small samples, EDMs typically predict 80%–90% of all outcomes and do better than what would be predicted by chance.

Overall, EDMs predict aggregate, group decision making, but with the appropriate ethnographic data, EDMs can also tell us about the rationales behind the 10%–20% of errors in a model. (More on this later.)

Ethnographic decision trees have a long history in the social sciences. Early EDMs were used to examine how farmers decided what to plant (Barlett 1980; Gladwin 1976, 1983, 1989b); how fish sellers set prices (Quinn 1978), how fisherman decided where to fish (Gatewood 1983); and how laypeople decided what to do when they or their children get sick (Mathews and Hill 1990; Young 1980). For more early examples, see Gladwin (1989a), Hill (1998), and Bernard et al. (2017).

The method of ethnographic decision tree modeling was made accessible in a landmark book by Christina Gladwin (1989a). Since then, the method has been applied to many kinds of choices. In addition to studies of illness decisions (Montbriand 1995; Ryan and Martínez 1996; Weller et al. 1997), researchers have used EDMs to explain: whether Navajo mothers breast-feed or use formula (Bauer and Wright 1996); when farmers in New Zealand use organic or conventional agricultural methods (Fairweather 1999); under what circumstances Houston's intravenous drug users decide to share needles or not (Johnson and Williams 1993); the decision to evacuate or not in the face of hurricanes (Gladwin et al. 2001); and how clinicians make outpatient referrals for patients with substance abuse (Breslin et al. 2000).

BK-TandF-ONWUEGBUZIE_9781138305267-200548-Chp16.indd 173 17/01/21 2:50 PM

More recently, EDMs have been used to describe whether or not to use a telecenter for entrepreneurial endeavors in Jamaica (Bailey and Ngwenyama 2013); how injured farmworkers make treatment timing choices (Thierry et al. 2015); whether women in rural Bangladesh give birth at a medical facility or at home (Edmonds 2010); and farmers' decisions in Uzbekistan to adopt agroforestry (Djalilov et al. 2016).

Technical Outline of EDMs for MMR: How to Build EDMs

Christina Gladwin (1989a) laid out five steps for building and testing EDMs: (1) select a specific behavioral choice to model; (2) elicit decision criteria from a purposive or convenience sample of respondents; (3) elaborate and verify the decision criteria on a purposive, heterogeneous sample of informants; (4) use the ethnographic data from Step 1 and the survey data from Step 2 to build a hierarchical decision model; and finally, (5) test the model on an independent and, if possible, representative sample from the same population. We (Ryan and Bernard 2006) added a sixth step in our study of people's decisions to recycle (or not) the last aluminum beverage can they had in their hand: (6) validate the model with responses from people about why they acted as they did. (For other clear descriptions of the steps, see Ryan and Martínez [1996], Hill [1998], and Beck [2005].)

Step 1 Selecting a Behavioral Choice to Investigate

The first step to building EMDs is for the researcher to identify the key decision to be explored and what alternatives are available. Fort (2011), for example, modeled whether people living in a protected forest reserve in southern Malawi decided to produce charcoal, an illegal livelihood activity; and Mwangi and Brown (2015) used the EDM approach to understand whether or not small and medium enterprises in Kenya decided to register for mobile banking services.

EDMs are not limited, however, to binary decisions. EDMs can also be used to predict choices among different alternatives. Oh and Park (2004) built models to predict whether cancer patients used hospital-only treatments or hospital-plus-alternative therapies. In a classic study, Young and Garro (1994) modeled how laypeople in a small rural town in Mexico decided among four treatment modalities: home treatments, treatment by folk curers (i.e., *curanderas*), treatment by local, unlicensed biomedical practitioners (i.e., *practicantes*), and treatment by a certified doctor.

Nor are EDMs limited to static descriptions. Following the lead of Young and Garro (1994), Ryan and Martínez (1996) modeled which of seven treatment alternatives (teas, carbonated beverages, rice water, sugar–salt solutions [SSS], pills, physical manipulations, and Western medical personnel) people living in San José, Mexico, used when their children had diarrhea. In addition to predicting whether a mother would use or not use each of treatment modalities, the Ryan–Martínez model also predicted which treatments mothers would use first, second, third, and fourth. In another example, Keshavarz and Karami (2014) built and tested three separate models to predict which of multiple coping mechanisms farmers in India would use when responding to initial, middle, and end stages of a prolonged drought.

Step 2 Eliciting the Decision Criteria

The second step is to identify the decision criteria that might influence individual choices. The objective is to elicit as many reasons as possible for why people make the choices that they do and to begin to understand *how* the criteria interact with each other to influence choices. Success in this step is determined by both the data elicitation techniques used as well as by who is selected to be interviewed.

There is no "correct" way to elicit decision criteria. Instead, researchers have used a variety of techniques, including informal interviews (Young and Garro 1994), in-depth ethnographic interviews (Shuk et al. 2012), more formal semi-structured interviews (Beck 2000), free listing tasks (Weller et al. 2016), paired comparisons (Keshavarz and Karami 2014 Young and Garro 1994), hypothetical scenarios (Roberts 2000), and participant observation (Bailey and Ngwenyama 2013; Beck 2000). Some elicitation protocols emphasize asking informants hypothetical questions such as "Why do people do or not do X?" while others ask about real past events, such as "The last time you had to choose between X and Y, what did you do and why?"

Most researchers use combinations of techniques to complement the strengths and offset the weaknesses of each particular method. For example, in Young and Garro's (1994) study in Mexico, they began with ethnographic fieldwork and had opportunities to observe how illness episodes unfolded. Then, in addition to informal talks and interviews, they used structured interview protocols including frame substitution, where they asked about the relationships between a set of propositions (Can you use X to cure A? Can you use X to cure B?) and a set of illnesses. They also used paired comparisons where they asked informants to compare pairs of treatment alternatives to each other (Is A better than B for curing X? For curing Y?).

From this combination of elicitation techniques, Young and Garro identified four criteria that consistently arose as important considerations in the choice of treatment: (1) illness gravity; (2) whether an appropriate home remedy was known; (3) faith in the effectiveness of the treatment for a given illness; and (4) expense of treatment and the availability of resources. For more on frame substitution and paired comparison techniques, see Weller and Romney (1988) and Borgatti (1994, 1999). For essential interviewing tips, see Spradley (1979), Becker (2008), and Bernard (2017).

To understand how Indian farmers coped with drought, Keshavarz and Karami (2014) conducted in-depth ethnographic interviews. They asked farmers to tell them how they had responded to the recent drought and recorded what the farmers did and the order in which they did it during the initial, middle, and end stages of the drought. To ensure that they didn't miss anything, they used a checklist of the drought management strategies they had compiled earlier to identify behaviors the farmers might have forgotten to mention. After collecting all the behaviors, they asked the farmers to explain why they decided to use particular strategies. They also used the paired comparison technique and asked under what circumstances the farmers would prefer one drought management strategy over another.

To elicit decision criteria from clinicians about whether to refer a patient to a substance abuse program, Breslin et al. (2000) conducted semi-structured interviews, but used a combination of hypothetical and real scenarios. They began by asking participants to explain in general the guidelines they use to make these choices. Next, they asked participants to think about a patient that they had recently referred to Program A and to describe (without identifying the patient) why they decided this particular program was appropriate. Then they asked participants to think about a patient that they had recently referred to Program B and to describe the factors that lead to that referral decision.

To identify how first-degree relatives of melanoma survivors made decisions about sun-protection behaviors, Shuk et al. (2012) conducted in-home, ethnographic interviews with 25 participants. In addition to observing the setting in which participants were making decisions, interviewers also asked participants to recall two separate recent sun-exposure periods when they were outdoors for an hour or more. For each sun-exposure event, they asked participants to report on whether they (1) used sunscreen; (2) sought shade; (3) put on a hat; and (4) used sun-protective clothing. They also asked participants to describe the event, including the outdoor activity and setting, who else was present, weather conditions, time of day, and the length of time they spent outdoors. Next, they asked informants for the reasons they used sun protection and reasons they did not use sun protection. Finally, they asked each person to answer a set of demographic questions. They used the information from the interviews to construct four meta decision-tree models (one for each sun-protection behavior).

Data collection at this point can also be broken into stages. Roberts (2000), for example, used a two-stage approach to elicit the factors that elder Jewish women in North Carolina who lived independently considered most significant when considering end-of-life decisions. Roberts first conducted unstructured and semi-structured qualitative interviews with a convenience sample of adults in their 50s. From these interviews, she identified six key conditions associated with end-of-life choices: (1) differing degrees of pain; (2) whether there was a cure for the disease; (3) reliance on others; (4) mental abilities; (5) financial burden worries; and (6) alternative living situations. She then used combinations of conditions to develop seven hypothetical scenarios that would be presented to older women. For instance, one scenario described a person living "in a nursing home with severe loss of intellectual functions and memory (i.e., Alzheimer's Disease) thus were totally reliant on others with no known cure with fairly-well controlled pain and no worry about being a financial burden" (Roberts 2000: 130).

In the second stage, Roberts conducted semi-structured interviews with a new sample of 13 Jewish women—nine women in their 50s, four women in their 70s, and one woman in her 90s. For each of the seven scenarios she had identified in the first stage, Roberts asked these 13 women whether they would (1) use aggressive treatment and life support measures; (2) not use aggressive treatment and life support measures (passive euthanasia); or (3) terminate one's life (active euthanasia). Roberts recorded each nominal response. Each scenario generated a lot of discussion (two–four hours), which Roberts audio-taped and transcribed verbatim. From these interviews, Roberts used the quantitative data from the responses to the scenarios, as well as the qualitative data from the discussions to identify pain and financial worries as being the most important conditions affecting whether women would choose euthanasia.

Empirical Demonstration of EDMs for MMR

For our study of whether to recycle cans, we asked 21 informants in North Dakota and Florida the following: (1) Think about the last time you had a can of something to drink in your hand-soda, juice, water, beer, whatever. When was that? (2) What did you do with the can when you were done? And (3) Why did you [didn't you] recycle? The first question generated a qualitative answer that we could dichotomize into "Recycled" or "Didn't Recycle." The second question generated a response that we could quantify into "Number of Days Ago." And the last question usually generated a short qualitative justification for their behavior. Recent literature on nonprobability sampling shows that 20–60 knowledgeable people are enough to uncover and understand core themes for a topic (Francis et al. 2010; Fugard and Potts 2015; Guest et al. 2006, 2017; Hagaman and Wutich 2017; Mason 2010; Morgan et al. 2002; Morse 1989; Sandelowski 1995; Weller et al. 2018).

We reached saturation (few new rationales mentioned for recycling or not). Table 16.1 shows the list of 27 reasons we retrieved for recycling and the 13 we retrieved for not cycling. Some people surely reported having recycled when they hadn't, but this doesn't affect the building of a preliminary model since good ethnographic informants know about the target behavior and can knowledgeably respond to questions about their own behavior (in this case, getting rid of an empty beverage can) and about their reasons for their behavior. The bottom line on sampling, then, is to maximize diversity (age, sex, ethnicity, education, area of the country, occupation) to elicit as wide a range of decision criteria as possible.

While convenience and purposive samples are common in the making of EDMs, Weller et al. (2016) were able to use the equivalent of a case-control comparison design to understand why some residents of Galveston, Texas, evacuated while others stayed behind during Hurricane Ike in 2008. Weller et al. began by identifying a sample of people from across the city who had evacuated. For each person interviewed, they matched them with a neighbor who did not evacuate. They chose the match-design so that "socioeconomic status and property damage would be distributed similarly in the two groups to allow a clearer focus on the rationale for evacuation." The researchers then conducted in-depth qualitative interviews, with open-ended questions, to elicit reasons, motives, and beliefs about evacuation, including why someone did or did not comply with evacuation orders, what they might do next time and why, and what they would like others to know who might be given an evacuation order in the future. They were careful to ask everyone the same questions. This allowed them to quantitatively compare the relative importance of all reasons in the decision-making process.

Step 3 Collecting Data for a Preliminary Model

The next step is to use the data from Step 2 to develop a structured elicitation instrument—one that asks everyone the same set of questions—and use these data to build a preliminary model of the behavior. Here again, yes/no questions produce data that make building a model easier. Ordering is also important. Ask about the behavior first—the thing you are trying to predict—then ask people to explain why they did what they did, and finally, ask about the conditions you wish to consider. In our (2006) study of recycling, we asked a purposive sample of 70 people 31 questions derived from Table 16.1. The 31 questions are shown in Table 16.2. Note that some of the questions generate dichotomous answers (e.g., yes, no); others generate nominal or qualitative response (e.g., at home, at work, etc.); and others generate quantitative responses (e.g., not at all, a little, some, a lot).

Why a sample of 70 at this stage? From our experience, we anticipated that our final decision model—pictured as a branching, bifurcating tree diagrams below, in Figure 16.1—would be at least three levels deep. To ensure that each of the decision's endpoints would contain at least five people, the minimum sample size would be $5*2^{(\#\ of\ Levels)}$, or $5*2^3 = 40$. Perfect bifurcation at each decision rarely happens, so we try to more-or-less double the minimum sample size to ensure that we wind up some cases at each endpoint, and hence, our 70 survey cases for building the model.

Some researchers use vignettes at this stage. Roberts (2000) wanted to model what an older Jewish woman would want to happen if she had an incurable illness and couldn't communicate. She recruited 91 women who identified as Reform Jews and 102 who identified as Conservative Jews, two of the main sects of American Judaism. Based on her initial elicitation interviews, Roberts developed ten hypothetical scenarios. For example, in the first scenario, Roberts asked, consider a situation where:

> You are in a coma and the doctors have said you are brain dead.
>
> You have no money left, all your financial resources have been exhausted and your only way to pay is through public benefits (Medicaid) and/or charity.
>
> Which choice would you make?
>
> 1. Use aggressive treatments and life supports to stay alive;
> 2. Not use aggressive treatments and no life supports to stay alive; or
> 3. Choose to terminate your life. (Roberts 2000: 41)

In the second scenario, Roberts asked: Consider a slightly different scenario where:

> You are in a coma and the doctors have said you are brain dead.
>
> There are enough funds to cover any of your needs, either through your own funds and/or your families' funds.
>
> Which choice would you make?
>
> 1. Use aggressive treatments and life supports to stay alive;

TABLE 16.1

Reasons for Recycling or not Recycling from 21 Informants

1. It's wasteful to just throw it away.	1. I was traveling and I had no place to recycle it.
2. The city has a recycling program. The garbage man picks it up.	2. Bins aren't around. I didn't have a recycling bin. There aren't enough recycling bins available.
3. To help save the environment.	3. There's no recycling program where I live. No city recycling program
4. Recycling bins are conveniently located.	4. Because I don't have big blue.
5. That's what big blue is for.	5. I didn't think about it.
6. My kid made a pact with a TV club so she now recycles.	6. I gave it to kids who turn it in for money.
7. I'm concerned about the environment.	7. Forgot.
8. It's environmentally sound.	8. Recycling is not available to me.
9. Land is not a renewable resource.	9. Laziness.
10. I save cans to get money for them.	10. The recycling bin was not conveniently located.
11. The people I'm staying with recycle, so I do, too.	11. Because I have to separate out cans from my garbage and that's a problem.
12. The bins were around.	12. Lack of education.
13. It's useful and can be used again.	13. I don't have enough time.
14. To keep the environment clean.	
15. Because of habit; we usually put it in big blue.	
16. Because I'm environmentally conscious.	
17. To preserve the environment for my kids.	
18. It's not biodegradable.	
19. It's no good in the landfill.	
20. Because it's just good to recycle.	
21. It's easy to do.	
22. Because it's the right thing to do.	
23. Because it's the big thing to do these days.	
24. Because someone told me to.	
25. We shouldn't cover the land up with garbage.	
26. To buy more beer.	
27. Because if you don't you have to pay a fee.	
28. Because if you don't you have to pay a fee.	

2. Not use aggressive treatments and no life supports to stay alive; or
3. Choose to terminate your life. (Roberts 2000: 41)

By examining how women's choices varied across the ten scenarios, Roberts concluded that in situations where a woman could no longer speak for herself and had little medical hope for recovery, pain was a greater influence than finances in end-of-life decisions.

Step 4 Building the Model

Figure 16.1 shows our preliminary model for recycling when people decide to recycle or throw away the last aluminum can they had in their hand. This step, described by Gladwin (1989a), takes a lot of trial and error. It involves taking each question from the survey in Step 3 and asking the following, for each one: If we could ask only this question, how many errors (that is, outcomes that are not predicted by the model) would result. For our study of whether people recycled the last aluminum beverage can they had in their hands, it turned out that question 7a in Table 16.2 (Were you at home at the time?) produced the fewest errors. That question became the first branch of our proposed model.

As you can see in Figure 16.1, guessing that everyone at home recycled and that everyone not at home did not recycle produces six errors (the four errors the model correctly predicts that people don't recycle and the two errors the model erroneously predicts that people will recycle) on the left-hand branch. On the right-hand branch, the model produces 11 errors (the 7 + 2 = 9 errors where the model correctly predicts people will recycle and the 1 + 1 = 2 errors where the model erroneously predicts that people don't recycle). Combined this is a total of 17 errors (6 + 11) and an accuracy rate of 53 out of 70 cases, or 76%. As you can also see, building models involves selecting different combinations of decision criteria (a qualitative process) and checking these combinations for accuracy (a quantitative process).

What this means is that by asking just one question (Were you home when you had that last can in your hand?) we can predict (for this sample) correctly 76% of the time. We try to improve on this by repeating the process and asking: What's the best predictor of recycling for each of the two branches of the model?

Of the 27 informants who were at home, the best predictor of who did or who did not recycle was to ask question 26 in Table 16.2: Do you recycle any materials besides cans? As you can see in Figure 16.1, of the 23 people who said they recycled other products, the model correctly predicts that 21 people (91.3%) recycled the last can they had in their hand and erroneously predicts that two people recycled when they didn't. And conversely, of the four people who said they didn't recycle other products, the model correctly predicts that all four (100%) didn't recycle. So, the rule here is this: For those at home who recycle other products, guess "recycled the can"; otherwise guess "didn't recycle the can." This results in just two errors out of 27 cases, or 92.6% correct.

On the right-hand branch of the model, just guessing that nobody recycled produced 32 out of 43 correct answers, or 74.4% correct. (Note, to calculate the total correct answers, we counted all the cases where the model correctly predicted that people didn't recycle (11 + 18 = 29), plus all the cases that the model erroneously predicted that people would recycle (1 + 2 = 3).) This improves to 88.4% correct by distinguishing whether those not at home were at work or elsewhere, and then asking: "Was a recycling bin conveniently located nearby?"

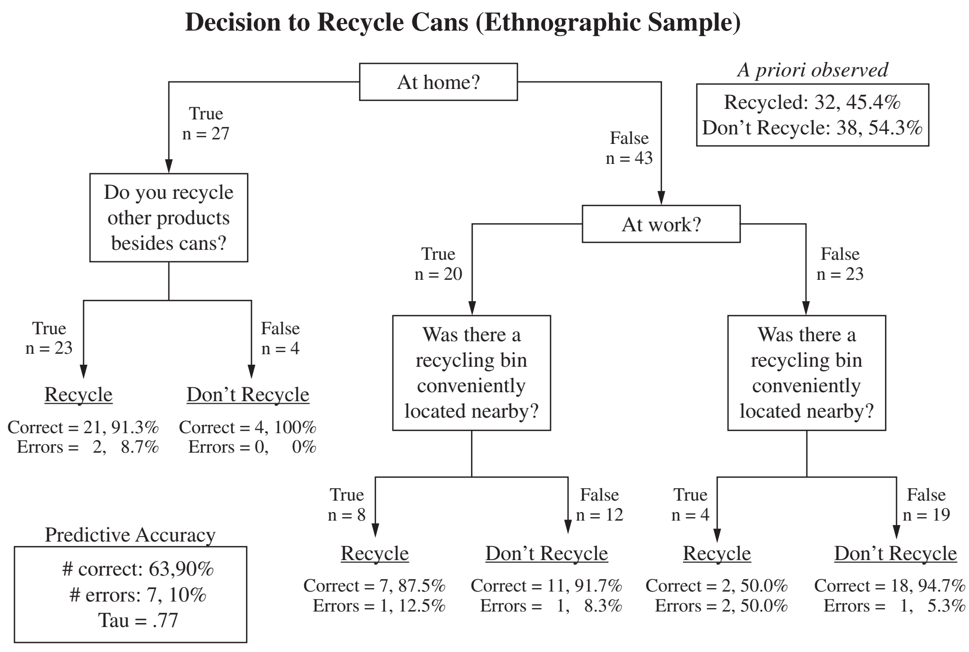

FIGURE 16.1 Decision to recycle cans (ethnographic sample, N=70)

TABLE 16.2

Questions Asked in the Recycling Study

After asking about the last can and the reasons for recycling, ask each of the following:

1. Does your city have a recycling program?
2. Can you return aluminum cans for redemption in your town or city?
3. Did you live in a house or apartment?
4. If you live in a house, is there a special pickup for recycled materials (e.g., big blue)?
5. Are there special bins for recycled materials in your apartment building etc.?
6. Are there recycling bins for cans where you work?
7. The last time you drank from an aluminum can were you:
   7a. at home?
   7b. at work?
   7c. driving in your car?
   7d. inside or outside?
   7e. at someone else's house?
8. The last time you drank from an aluminum can did you get the can from a vending machine?
9. The last time you drank from an aluminum can was there a recycling bin conveniently located nearby?
10. The last time you drank from an aluminum can were you busy?
11. The last time you drank from an aluminum can were there other people around when you finished your drink?
12. If so, do these people usually recycle cans?
13. If so, did anyone suggest that you recycle the can?
14. Do you have children?
15. Do you habitually recycle material such as cans, newspapers, and plastics at home?
16. Do you habitually recycle material such as cans, newspapers, and plastics at work?
17. Do you consider yourself environmentally conscious (not at all, a little, some, a lot)?
18. How much do you think that recycling helps to save the environment (not at all, a little, some, a lot)?
19. How much are you concerned about the environment (not at all, a little, some, a lot)?
20. How much do you think recycling helps to keep the environment clean (not at all, a little, some, a lot)?
21. How important is it for you to preserve the environment for children (not at all, a little, some, a lot)?
22. Do you think it's wasteful to throw away an aluminum can?
23. Do you think that there is a lot of social pressure nowadays to recycle?
24. Do you think that cans are bad for landfills?
25. Do you think that recycling aluminum cans is useful?
26. Do you recycle any materials besides cans?
27. If so, what other materials do you recycle?

Note: Questions 1–16 and questions 22–26 were answered Yes or No. Questions 7–11 are expansions of the question "Where were you when you had that used beverage can in your hand?" into five binary questions. This ensures that all informants are given the same set of cues as the data are collected to build the preliminary model. Questions 17–21 produced ordinal data (not at all, a little, some, a lot). Question 27 produced text.

First, as shown in Figure 16.1, when people were at work and bins were nearby, seven of the eight respondents who were at work recycled (the eight includes the seven people who the model correctly predicted would recycle and the one person the model erroneously predicted didn't recycle). When people were at work and bins were unavailable, 11 of the 12 respondents said that they didn't recycle (the 12 includes the 11 people who the model correctly predicted didn't recycle and the one person the model erroneously predicted didn't recycle). This branch of the model gets 18 out of 20, or 90% correct. Second, among the 23 respondents who were neither at home nor at work, asking if a recycling bin was nearby produces a model with just three errors, or 87% correct. Overall, on the right-hand branch, the model produces five errors (88.4% correct), and the accuracy of the complete model (both left- and right-hand branches) is 63 right out of 70, or 90%.

So overall, how did we do? In Figure 16.1, the box labeled "a priori observed," refers to the how many people reported recycling (32 out of 70 or 45.7%) and how many people reported not recycling (38 out of 70 or 54.3%). If this was all we knew and we had to guess whether a can gets recycled or not, then, without any information about the circumstances (being at home, having a recycling bin nearby, etc.), our best guess on any individual case would have been that people didn't recycle and we would have been correct 54.3% of the time and wrong 45.7% of the time. We refer to this as "a priori" because this tells us how accurate we would be before we built our model. But if we use the model described here, then we do much better. In fact, the model predicts 77% better than expected by chance (and is indicated by Klecka's tau = .77 in Figure 16.1. (See Klecka 1980: 50–51 for details on how to calculate this statistic).

As it happens, if the only criterion for success is reducing the number of errors to a minimum, the model in Figure 16.1 is actually more complicated than it needs to be. As Figure 16.2 shows, we can collapse the two paths "At Work?" and "Not at Work?" without any loss of predictive power—it's the same 90% for the models in both Figure 16.1 and Figure 16.2. The reason is this: Most of the predictive power on the right side of the model is based on a bin being nearby. The extra criterion in the model, however (at work–not at work in Figure 16.1), with its two extra paths, shows that people at work recycle more than do those who are neither at home nor at work—40% (7 + 1 = 8 of 20), compared to 13% (2 + 1 = 3 of 23). The extra criterion thus provides information on the size and location of the problem—information that suggests where to put recycling bins if we don't have an unlimited supply of them.

Step 5 Testing the Model on an Independent Sample

An accuracy rate of 90% may seem high, but it's hardly a surprise when a model accounts for the data on which it is built. All we are in doing in Figure 16.1 is representing graphically what people told us they did. Models, however, are tentative theories, or hypotheses. Their validity doesn't depend on how they are derived but on how well they stand up to tests on an independent sample of people who were not involved in building the model in the first place. Model confirmation on independent samples is typically done in EDM studies to increase their internal and external validity.

Strong agreement between two or more, independently derived EDMs is similar to repeating a laboratory experiment in terms of reliability and internal validity. For example, assume another group of independent researchers followed the process describe above but collected data from different people at each step. If after all the steps and decisions the

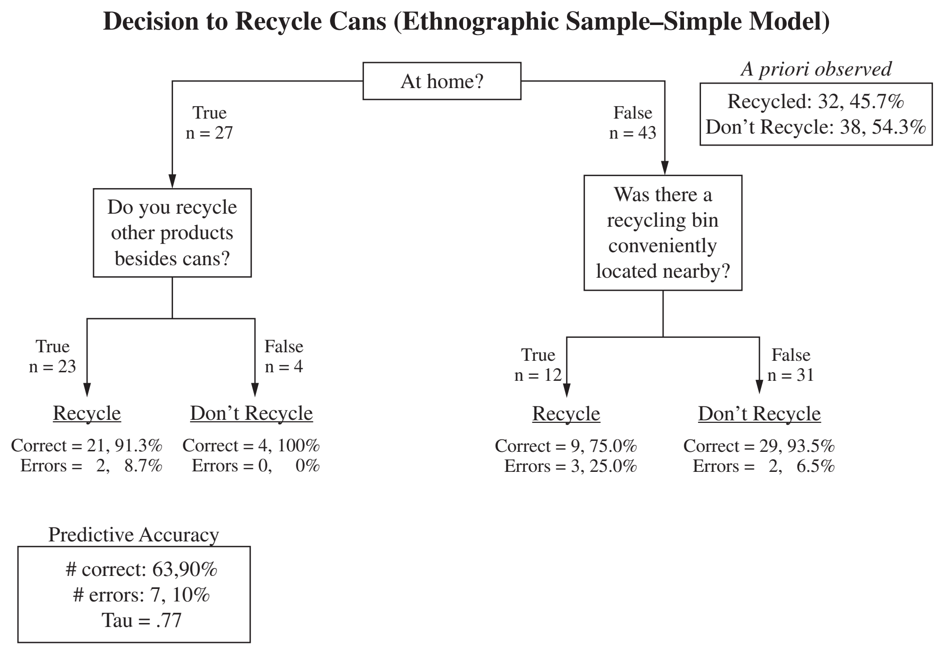

FIGURE 16.2 Decision to recycle cans—simplified model (ethnographic sample, N=70)

other research team made, the prediction based on their data resulted in a model that resembled ours, we would have a lot of confidence that the results were relatively robust and unlikely to be spurious. Still, as with all ethnographically derived findings, there is doubt about external validity—whether the results can be generalized to a larger population (Weller et al. 1997). This is just another reason why it is so important to collect detailed information about the characteristics of your participants.

To see how our model EDM did on a larger population, we tested it on a representative, national phone survey of 386 respondents in the United States. The results are in Figure 16.3.

Comparing Figures 16.1 and 16.3, there are many differences in the distribution of answers. For the ethnographic model, we interviewed people wherever we could find them, while for the national survey, we called people at home. Sure enough, 58% of the national respondents said they were at home when they had that last beverage can in their hand, compared to only 39% of our ethnographic informants.

And still, the ethnographic model held up well in a national test. Of the 173 people who were at home and who also said they recycled other products besides cans, 160 (93%) recalled recycling the can, compared to 91% in the ethnographic sample. Of the 55 people at home who reported not recycling other products besides cans, 45 (82%) recalled not recycling the can, compared to 100% for the ethnographic sample. And the overall accuracy of the national model is 85%, compared to 90% for the ethnographic model.

What this means is that there is likely to be a strong consensus, across the United States, about the decision to recycle aluminum beverage cans. We had a glimpse of this when we got the same model in North Dakota and Florida, but of course, it didn't have to turn out that way. Multiple ethnographically derived models on small samples could easily turn up cultural or regional differences for any given decision—like when to take a sick child to the doctor or when to report child abuse.

Just as with the ethnographic model, adding the question in the national model about where the behavior took place (at work vs. at home) had no effect on prediction power. It did, however, corroborate the policy-relevant information produced in the ethnographic model regarding where to put scarce resources if we want to increase recycling behavior. Of the 158 people in the national sample who said they were not at home when they had that last beverage can in their hands, 20% (20 + 1 = 21 out of 104 in Figure 16.3) said they didn't recycle if they also said they were at work.

By contrast, 50% (25 + 2 = 27 out of 54 in Figure 16.3) of the not-home people said they didn't recycle if they also said they were not at work. It may be tempting to try to reduce the 50% error rate by further modifying the model, but the not-home/not-at-work condition covers people at who are football games, or driving on the freeway, or visiting other people's houses, or window shopping. With so many conditions, and limited resources with which to put out recycling bins, it is going to be tough to have an impact on that 50% error rate. In the short term, it's easier to imagine incentives for getting employers to put out those bins.

Step 6 Assessing the Validity of Ethnographic Decision Models

Thus far, we have focused on the quantitative measure of model error rate. We can take the process one step further and bring qualitative information into assessing the model's validity.

We asked 33 of our ethnographic informants to tell us, in their own words, why they recycled, before we asked systematically about the decision-criteria. For these informants, we examined the fit between their justifications of their choice to recycle or not and the model's predictions. We did this by

BK-TandF-ONWUEGBUZIE_9781138305267-200548-Chp16.indd 179 17/01/21 2:50 PM

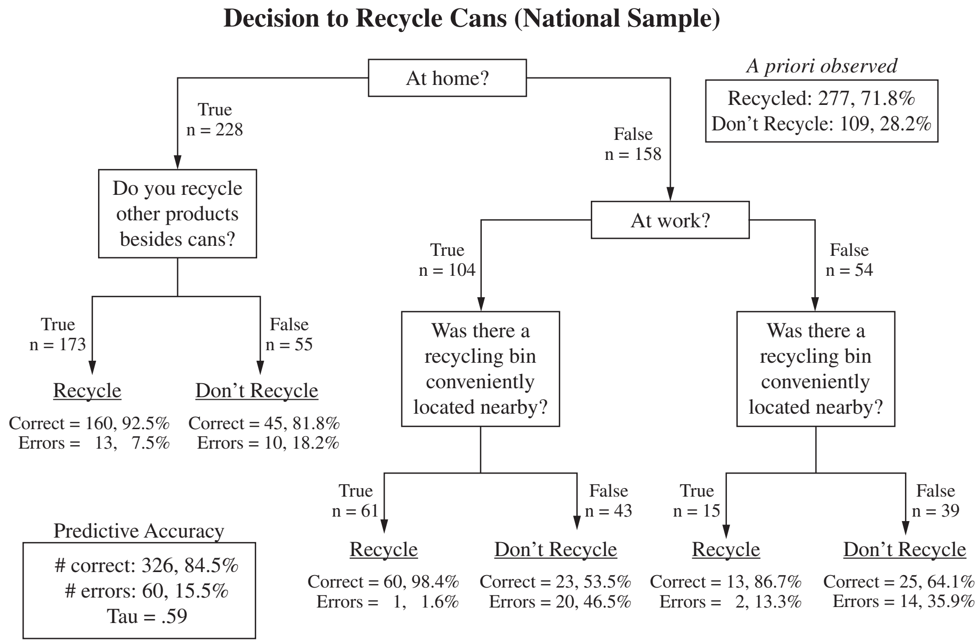

FIGURE 16.3 Decision to recycle cans (national sample, N=386)

following individual recycling cases down the decision tree and examining the degree to which each end point in the tree (each final decision) corresponded to our informants' own accounts. The results are shown in Table 16.3.

The top right-most cell of Table 16.3 contains the rationales from people who reported that they were at home and recycled other things besides cans. At home, only one person said it was easy to do. Three people, however, said it was a good thing to do or that it was mandatory. Nowhere else in the rationales do these latter two themes arise.

The next cell down shows the rationales from people who reported that they were at home but did not recycle other things. The model correctly predicted that the first three of the respondents would not recycle. Unlike those who had recycled, none of the three mentioned that it was important or good to recycle-nor that it was mandatory. The three cases that were misclassified more closely resemble the rationales in the cell above.

The rationales for those at work are clearly divided between those respondents who reported having a recycling bin conveniently located nearby and those who did not. Those who had a bin nearby reported its availability and the ease with which one could recycle. Those who didn't have a bin spontaneously mentioned not being able to keep cans on the job or not having a recycling center.

The last cell shows the rationales for those who were not at home or at work and who did not have a recycling bin nearby. Of the eight cases that the model predicted correctly, half spontaneously mentioned either the lack of convenience or the lack of a recycling bin. The other half mentioned explicitly where they were and clearly implied that place had something to do with their behavior. The two people who said that they threw the can out of the car identified a factor we hadn't thought of before—that laws against drinking and driving might have an impact on environmentally friendly behaviors.

Resolving Errors with Ethnography

Why do some people recycle when the model predicts that they shouldn't and other people don't recycle when the model predicts that they should? We turn to the verbatim comments of our informants. Those who recalled recycling a can despite not being at home or at work and not having a bin conveniently located were likely to justify their behavior by citing their beliefs in environmentalism or citing financial benefits for doing it. This may be the result of positive attitudes about recycling—attitudes that give people the extra impetus they need for recycling when bins aren't handy.

This is worth testing, but note that attitudes (or whatever else is at work) can account for no more than 10% of responses in the local sample (since the model predicts 90% of responses there) and no more than 15% of responses in the national sample (since the model predicts 85% of responses there).

We don't have ethnographic data to account for those not at home who reported not recycling a can despite having a recycling bin handy because none of our 33 informants were in that category. In fact, only three out of 76 people in our national sample reported not recycling despite having a bin handy. Just putting a lot of recycling bins around will likely increase recycling behavior. This has been known for some time, of course, but the fact that we can validate a well-understood piece of information like this gives us confidence in the EDM method for answering questions that are not this obvious.

Suggested Applications of EDMs in MMR

In principle, ethnographic models can be used to study complex choices, but in practice they are most commonly used for binary (yes/no) or categorical (A, B, or C) decisions. Binary decisions can, of course, be modeled statistically,

BK-TandF-ONWUEGBUZIE_9781138305267-200548-Chp16.indd 180 17/01/21 2:50 PM

TABLE 16.3

Verbatim Justifications for 33 Recycling Choices from Ethnographic Sample

Decision Rules	Choice	Verbatim Justification
At home? Yes Recycle other things? Yes	Yes	I know you can recycle it and the bin was easy to get to. I believe in it, and it's good for the environment. I feel that it's some form of token effort in trying to protect the environment and keep stuff out of landfills. I recycle as much as I can. For recycling—because garbage just doesn't disappear—if you recycle there is less garbage then. They pick it up on Wednesday—because it's a good thing to do. It is required to recycle cans. It is mandatory, and I believe in recycling. It is mandatory.
At home? Yes Recycle other things? No	No	*Correct* I don't recycle. I didn't think about it and I don't like storing it around home because it brings pests. I was too lazy. Sometimes I keep em' for my brother but … I give them to him.… I just didn't this time. *Incorrect* I take them to a place where they take aluminum cans and gets money for em'. I did it to recycle … no reason just to recycle. It's easy to do and they pick em' up.
At work? Yes Bin nearby? Yes	Yes	I always recycle aluminum cans. … I don't know … because I can, because it's available. It's an automatic thing at work; we all recycle there. I wasn't gonna mess with it—it was easy. One of the operator collects them at work, and she takes the bag weekly to put it … to take in for recycling.
At work? Yes Bin nearby? No	No	We're not allowed to keep cans on the job. There was no recycling center nearby. A lady at work collects them—so I put them in the bag to give to this one lady.
At work? Yes Bin nearby? Yes	No	*Correct* It wasn't convenient I guess. There was no obvious place to put it for recycling. I don't know—I didn't have a container to put it in. I was not home—I was someplace in town. I was at someone's house. I was driving—I threw it out the window—it was a beer can—the environment—I'm down with it but there are too many rules—I threw it out so I wouldn't get caught with it in my car. I wasn't at home—at home I would've put it in the recycling bucket — If it weren't illegal to put it in my car.… I'd've taken it home with me—more people would recycle if it weren't for those open container laws. *Incorrect* Well I didn't know what to do with it. That's better on the environment. I take em' in and turns em' in for money. I think it's a good thing—why use new things when you can reuse old things.

with logistic regression and categorical decisions can be modeled statistically with multinomial logit models or conditional logit models, but *ethnographic* decision models (as contrasted with purely statistical ones) are particularly good for understanding real-life, discrete choices and the rationales behind those choices. Unlike most purely probabilistic models, EDMs depend on a series of related if-then choices—things that we think make a lot of sense in real-world settings.

Strengths and Limitations of EDMs in MMR

EDMs are based on mixed methods during the full inquiry process—from exploratory data collection and analysis, to initial model building and hypothesis generation, to confirmatory analysis for model accuracy and validation. EDMs involve fully inductive work—inducing rules from a set of real-life events—and fully deductive work—generating and confirming hypotheses. On the other hand, while EDMs are typically generated on small samples and have high internal validity, they require testing on large samples in order to check external validity. Without this second step, EDMs are not likely to be convincing in applications research, where the goal is to get people to change their behavior—to recycle beverage cans, for example, as in the case reported in this chapter.

Resources for Learning More about EDMs

Gladwin (1989a) remains the key source for anyone who want to learn more about EDMs. Other overviews include Beck 2005, Bernard et al. (2017), and Hill (1998). See Ryan and Martínez (1996) and Weller et al. (1997) for more nuances in regard to accuracy and predictability.

REFERENCES

Bailey, A., and O. Ngwenyama. 2013. Toward entrepreneurial behavior in underserved communities: An ethnographic decision tree model of telecenter usage. *Information Technology for Development* 19: 230–248. doi:10.1080/02681102.2012.751571

Barlett, P. F. 1980. Cost-benefit analysis: A test of alternative methodologies. In *Agricultural decision making*, edited by P. Barlett, 137–160. New York: Academic Press.

Bauer, M., and A. Wright. 1996. Integrating qualitative and quantitative methods to model infant feeding behavior among Navajo mothers. *Human Organization* 55: 183–192. doi:10.17730/humo.55.2.p55g316v70572732

Beck, K. A. 2000. A decision making model of child abuse reporting. Doctoral dissertation, Vancouver, University of British Columbia. Retrieved from https://open.library.ubc.ca/cIRcle/collections/ubctheses/831/items/1.0089769

Beck, K. A. 2005. Ethnographic decision tree modeling: A research method for counseling psychology. *Journal of Counseling Psychology* 52: 243. doi:10.1037/0022-0167.52.2.243

Becker, H. S. 2008. *Tricks of the trade: How to think about your research while you're doing it*. Chicago: University of Chicago Press.

Bernard, H. R. 2017. *Research methods in anthropology*, 6th ed. Lanham, MD: Rowman & Littlefield.

Bernard, H. R., A. Wutich, and G. W. Ryan. 2017. *Analyzing qualitative data: Systematic approaches* (2nd ed.), Thousand Oaks, CA: Sage.

Borgatti, S. P. 1994. Cultural domain analysis. *Journal of Quantitative Anthropology* 4: 261–278. Retrieved from http://works.bepress.com/steveborgatti/44/

Borgatti, S. P. 1999. Elicitation techniques for cultural domain analysis. *Enhanced Ethnographic Methods*3: 115–151.

Breslin, F. C., C. H. Gladwin, D. Borsoi, and J. A. Cunningham. 2000. Defacto client-treatment matching: How clinicians make referrals to outpatient treatments for substance use. *Evaluation and Program Planning* 23: 281–291. doi:10.1016/S0149-7189(00)00014-8

Corbin, J., and Strauss, A. L. 2015. *Basics of qualitative research. Techniques and procedures for developing grounded theory* (4th ed.). Thousand Oaks, CA: Sage.

Djalilov, B. M., A. Khamzina, A. K. Hornidge, and J. P. Lamers. 2016. Exploring constraints and incentives for the adoption of agroforestry practices on degraded cropland in Uzbekistan. *Journal of Environmental Planning and Management* 59: 142–162. doi:10.1080/09640568.2014.996283

Edmonds, J. K. 2010. Social networks, decision making and use of skilled birth attendants to prevent maternal mortality in Matlab, Bangladesh. Doctoral dissertation, Atlanta, Emory University. Retrieved from https://etd.library.emory.edu/

Fairweather, J. R. 1999. Understanding how farmers choose between organic and conventional production: Results from New Zealand and policy implications. *Agriculture and Human Values* 16: 51–63. Retrieved from https://link.springer.com/journal/10460

Fort, J. D. 2011. The role of factors involving the environment in a forest livelihood decision of Malawian villagers. Master's thesis, Gainesville, University of Florida. Retrieved from http://ufdc.ufl.edu/UFE0043443/00001

Francis, J. J., M. Johnston, C. Robertson, L. Glidewell, V. Entwistle, M. P. Eccles, and J. M. Grimshaw. 2010. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology and Health* 25: 1229–1245. doi:10.1080/08870440903194015

Fugard, A. J., and H. W. Potts. 2015. Supporting thinking on sample sizes for thematic analyses: A quantitative tool. *International Journal of Social Research Methodology* 18: 669–684. doi:10.1080/13645579.2015.1005453

Gatewood, J. B. 1983. Deciding where to fish: The skipper's dilemma in southeast Alaskan salmon seining. *Coastal Management* 10: 347–67. doi:10.1080/08920758309361928

Gladwin, C. H. 1976. A view of the Plan Puebla: An application of hierarchical decision models. *American Journal of Agricultural Economics* 58: 881–887. Retrieved from https://academic.oup.com/ajae

Gladwin, C. H. 1983. Contributions of decision-tree methodology to a farming systems program. *Human Organization* 42: 146–57. Retrieved from www.sfaa.net/publications/human-organization/

Gladwin, C. H. 1989a. *Ethnographic decision tree modeling.* Newbury Park, CA: Sage.

Gladwin, C. H. 1989b. Modeling farmers' decisions to change: Using cognitive science in the design of agricultural technology. *Social science perspectives on managing agricultural technology*, edited by D. Groenfeldt and J. W. Mook, 127–141. IWMI Books, International Water Management Institute, number 113863.

Gladwin, C. H., H. Gladwin, and W. G. Peacock. 2001. Modeling hurricane evacuation decisions with ethnographic methods. *International Journal of Mass Emergencies and Disasters* 19: 117–143. Retrieved from http://ijmed.org/

Glaser, B., and A. Strauss. 1967. *The discovery of grounded theory. Strategies for qualitative research*. Chicago: Aldine.

Guest, G., A. Bunce, and L. Johnson. 2006. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 18: 59–82. doi:10.1177/1525822X05279903

Guest, G., E. Namey, and K. McKenna. 2017. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field Methods* 29: 3–22. doi:10.1177/1525822X16639015

Hagaman, A., and A. Wutich. 2017. How many interviews are enough to identify metathemes in multi-sited and cross-cultural research? Another perspective on Guest, Bunce, and Johnson's (2006) landmark study. *Field Methods* 29: 23–41. doi:10.1177/1525822X16640447

Hill, C. E. 1998. Decision modeling: Its use in medical anthropology. In *Using methods in the field*, edited by V. C. de Munck and E. J. Sobo, 139–64. Walnut Creek, CA: AltaMira Press.

Johnson, J., and M. L. Williams. 1993. A preliminary ethnographic decision tree model of injection drug users' (IDUs) needle sharing. *International Journal of the Addictions* 28: 997–1014. Retrieved from www.tandfonline.com/toc/isum19/25/sup1?nav=tocList

Keshavarz, M., and E. Karami. 2014. Farmers' decision-making process under drought. *Journal of Arid Environments* 108: 43–56. doi:10.1016/j.jaridenv.2014.03.006

Klecka, W. R. 1980. *Discriminant analysis.* Beverly Hills, CA: Sage

Krippendorf, K. 2013. *Content analysis* (3rd ed.). London: Sage.

Mason M. 2010. Sample size and saturation in PhD studies using qualitative interviews. *Forum: Qualitative Social Research*, 2010; 11. doi:10.17169/fqs-11.3.1428

Mathews, H. F., and C. Hill. 1990. Applying cognitive decision theory to the study of regional patterns of illness treatment choice. *American Anthropologist* 91: 155–170. doi:10.1525/aa.1990.92.1.02a00110

Montbriand, M. J. 1995. Decision tree model describing alternate health care choices made by oncology patients. *Cancer Nursing* 18: 117. Retrieved from https://journals.lww.com/cancernursingonline/pages/default.aspx

Morgan, M. G., B. Fischoff, A. Bostrom, and C. J. Atman. 2002. *Risk communication: A mental models approach*. New York: Cambridge University Press.

Morse, J. M. 1989. Strategies for sampling. In *Qualitative nursing research: A contemporary dialogue*, edited by J. M. Morse, 117–31. Rockville, MD: Aspen Press.

Murthy, S. K. 1998. Automatic construction of decision trees from data: A multi-disciplinary survey. *Data Mining and Knowledge Discovery* 2: 345–389. Retrieved from https://link.springer.com/journal/10618

Mwangi, B. J., and I. Brown. 2015. A decision model of Kenyan SMEs' consumer choice behavior in relation to registration for a mobile banking service: A contextual perspective. *Information Technology for Development* 21: 229–252. doi:10.1080/02681102.2013.874320

Oh, H. S., and H. A. Park. 2004. Decision tree model of the treatment-seeking behaviors among Korean cancer patients. *Cancer Nursing* 27: 259–266. Retrieved from https://journals.lww.com/cancernursingonline/pages/default.aspx

Quinn, N. 1978. Do Mfantse fish sellers estimate probabilities in their heads? *American Ethnologist* 5: 206–226. doi:10.1525/ae.1978.5.2.02a00020

Roberts, E. 2000. Women's choices for end of life care: How older Jewish women make clinical decisions. Doctoral dissertation, Columbia, University of South Carolina. Retrieved from https://search.proquest.com/pqdtglobal/index?accountid=10920

Ryan, G. W., and H. R. Bernard. 2006. Testing an ethnographic decision tree model on a national sample: Recycling beverage cans. *Human Organization* 65: 103–114. doi:10.17730/humo.65.1.884p8d1a2hxxnk79

Ryan, G. W., and H. Martinez. 1996. Can we predict what mothers do? Modeling childhood diarrhea in rural Mexico. *Human Organization* 55: 47–57. doi:10.17730/humo.55.1.p0n23832j8q34743

Sandelowski, M. 1995. Sample size in qualitative research. *Research in Nursing and Health* 18: 179–183. doi/pdf/10.1002/nur.4770180211

Shuk, E., J. E. Burkhalter, C. F. Baguer, S. M. Holland, A. Pinkhasik, M. S. Brady, and J. L. Hay. 2012. Factors associated with inconsistent sun protection in first-degree relatives of melanoma survivors. *Qualitative Health Research* 22: 934–945. doi:10.1177/1049732312443426

Spradley, J. 1979. *The ethnographic interview*. New York: Holt, Rinehart and Winston.

Thierry, A. D., and S. A. Snipes. 2015. Why do farmworkers delay treatment after debilitating injuries? Thematic analysis explains if, when, and why farmworkers were treated for injuries. *American Journal of Industrial Medicine* 58: 178–192. doi:10.1002/ajim.22380

Weller, S. C., R. Baer, and J. Prochaska. 2016. Should I stay or should I go? Response to the Hurricane Ike Evacuation order on the Texas Gulf Coast. *Natural Hazards Review* 17: 04016003. doi:10.1061/(ASCE)NH.1527-6996.0000217

Weller, S. C., and A. K. Romney. 1988. *Systematic data collection*. Newbury Park: Sage.

Weller, S. C., T. R. Ruebush, and R. E. Klein. 1997. Predicting treatment-seeking behavior in Guatemala: A comparison of the health services research and decision–theoretic approaches. *Medical Anthropology Quarterly* 11: 224–45. doi:10.1525/maq. 1997.11.2.224

Weller, S. C., B. Vickers, H. R. Bernard, A. M. Blackburn, S. Borgatti, C. C. Gravlee, and J. C. Johnson. 2018. Open-ended interview questions and saturation. *PLoS ONE* 13(6): e0198606. doi:10.1371/journal.pone.0198606

Young, J. C. 1980. A model of illness treatment decisions in a Tarascan town. *American Ethnologist* 7: 106–131. doi:10.1525/ae.1980.7.1.02a00070

Young, J. C., and L. C. Garro. 1994. *Medical choice in a Mexican village*. New Brunswick, NJ: Rutgers University Press.

Zhang, Z. 2016. Decision tree modeling using R. *Annals of Translational Medicine* 4: 275. doi.org/10.21037/atm.2016.05.14