Elicitation Techniques: Getting People to Talk About Ideas They Don’t Usually Talk About

Keith C. Barton

Indiana University

Published online: 26 May 2015.

To cite this article: Keith C. Barton (2015) Elicitation Techniques: Getting People to Talk About Ideas They Don’t Usually Talk About, Theory & Research in Social Education, 43:2, 179-205, DOI: 10.1080/00933104.2015.1034392

To link to this article: http://dx.doi.org/10.1080/00933104.2015.1034392
Elicitation Techniques: Getting People to Talk About Ideas They Don’t Usually Talk About

Keith C. Barton
Indiana University

Abstract: Elicitation techniques are a category of research tasks that use visual, verbal, or written stimuli to encourage participants to talk about their ideas. These tasks are particularly useful for exploring topics that may be difficult to discuss in formal interviews, such as those that involve sensitive issues or rely on tacit knowledge. Elicitation techniques can also reduce power imbalances between interviewers and respondents, and they can enhance participants’ ability to elaborate on their own conceptions of the world, rather than limiting them to categories derived from theory or previous research. Among the most useful of such techniques are those that involve respondents in arranging stimulus materials, constructing materials in response to stimuli, and explaining stimulus materials. Each of these has been used to explore important topics in social education, and familiarity with a range of elicitation techniques enables researchers to overcome many barriers to productive interviewing.

Keywords: elicitation techniques, interview research, research methods

Social studies researchers often ask about topics that are difficult to discuss in formal interviews. These include controversial political topics; abstract concepts, such as democracy or human rights; issues that touch on ethnic, religious, gender, or sexual identities; ideas that may be largely tacit, such as those related to teaching and schooling; and conceptual understandings that are rarely discussed in everyday conversation (e.g., historical time, significance, or evidence). Asking straightforward questions about these subjects can be frustrating, because many participants either cannot talk about them easily or prefer not to, even with a trusted researcher—much less with someone they barely know. However, a class of research tasks known as elicitation techniques...
can facilitate such conversations by displacing the focus of interviews onto external stimuli and, in some cases, changing the power balance between researchers and participants. Familiarity with a range of such techniques can help researchers collect rich data even on difficult topics.

Many researchers, however, are not aware of the wide variety of elicitation techniques available to them. In part, this may be due to the diverse origins of these tasks, many of which originated in sociology, anthropology, or psychology. Commonly used texts in educational research, meanwhile, give scant attention to elicitation techniques (e.g., Creswell, 2008; Fraenkel & Wallen, 2006; Patton, 2002), and even standard works on interview research devote little or no space to the topic (e.g., Brinkmann & Kvale, 2015; Seidman, 2013). As a result, even though elicitation techniques are common in educational research, researchers may only be familiar with the use of particular tasks in specialized areas of study, such as concept maps in science education or stimulated recall in studies of teachers’ decision making. If researchers have only encountered one such method, they may not realize that they can choose from a range of options, each suited to a different purpose. The range of tasks may also be obscured by a lack of consistent terminology or shared theoretical assumptions: Sometimes these tasks (or a subset of them) are referred to as *semi-projective devices* (Binford, 1984), *visual methods* (Harper, 1994), or *structured interviewing methods* (Bernard, 2006), or as part of approaches such as phenomenography (Marton, 1986), Piagetian clinical interviewing (Ginsburg, 1997), or personal construct psychology (Ryle, 1975). This specialized and varied terminology, combined with a lack of elaboration in research methods texts, makes it difficult to see the wider landscape of elicitation. This article aims not to provide a definitive account of elicitation techniques but to chart some of the contours of the field so that researchers can better see the variety of paths available to them.

**OVERCOMING BARRIERS TO DISCUSSION**

Although elicitation techniques have no single or authoritative definition, the concept most often refers to research tasks that use visual, verbal, or written stimuli to encourage people to share their ideas (e.g., Johnson & Weller, 2002). Such methods provide useful alternatives to direct questions about participants’ thinking, particularly when social, cultural, or psychological barriers make it difficult to talk about a topic. Although both children and adults can easily articulate their ideas about some topics, they have much less experience talking about others. Ask primary teachers how they teach the number line, and they will probably be able to answer in detail. Ask 8-year-olds how their classmates behave on the bus and they will provide rich data. But ask the same teachers how they provide emotional support to students, or the 8-year-olds how social life has changed over the last two centuries, and responses will be much less...
Elicitation Techniques

complex. Researchers make their living with words, and it is hardly surprising they consider open-ended, verbal interviews a comfortable methodology. But not all participants have either the background or the inclination to engage in extended, abstract verbal exchanges.

A number of factors affect participants’ willingness and ability to discuss their ideas during formal interviews. As Tobin, Hsueh, and Karasawa (2009) point out, teachers can find it difficult to answer a question such as, “What is your philosophy of classroom management?” both because it is too abstract and because it sounds too much like an exam question. As a result, asking broad, open-ended questions or soliciting the kind of “grand tour” responses recommended in ethnographic interviews (Spradley, 1979) is often ineffective. In fact, such nondirective interviews actually can create anxiety among respondents, because they may not know what is “really” being asked or what expectations the researcher brings to the setting (Whyte, 1984). Children (and others who consider themselves less knowledgeable than researchers) may feel that they are being tested on their ability to supply the right answers. The length and complexity of participants’ responses are also affected by their memory, their perception of the relevance of questions, and their own feelings of self-worth or social desirability (Johnson & Weller, 2002). Controversial topics in particular may lead respondents to provide answers that they feel will make a positive impression on the researcher, while topics they perceive as taboo may lead them to offer little more than vague generalities.

Lack of elaboration, however, does not necessarily indicate lack of knowledge, nor does it signal mild opinions or lack of interest. Participants may know a great deal about a topic and have a great deal to say, but special tasks are sometimes necessary to bring their ideas to the surface, as well as to encourage them to articulate those ideas in deeper and more complex ways. Elicitation techniques accomplish this by making the research process more transparent, comfortable, and authentic. Unlike written instruments, classroom observations, or less structured interviews, elicitation techniques allow participants to more easily see what researchers hope to accomplish. Walker and Widel (1985) tell of a teacher who was engaged both in their own project with visual stimuli and another more traditional research study. The teacher noted that despite the other researcher’s explanations, she never really understood the study or felt a part of it. “Somehow,” she said, “it all seems to disappear into his notebook and never come back.” But she noted that in the elicitation study, “I can see what you are doing. I can get interested it, and so can the kids” (p. 213).

In addition to their transparency, elicitation tasks can make the interview setting more comfortable (Bernard, 2006; Catterall & Ibbotson, 2000; Whyte, 1984). Putting items in order, commenting on photographs, or sketching a diagram are activities that provoke little anxiety. Even young children are usually not afraid they will fail to live up to researchers’ expectations. Many such tasks mirror conversations people have in daily life, such as commenting on the images in a family photo album (Schwartz, 1989). Looking
at photographs can cover otherwise awkward silences, and there is no need for direct eye contact. Participants even become less self-conscious about note-taking or recording equipment because questions focus on materials rather than the respondent. Even when researchers probe participants’ responses, the presence of physical materials—particularly visual images—can make this process less threatening. Asking and answering a set of purely verbal questions, especially when researchers and participants come from different backgrounds, seems a lot like schooling, and respondents sometimes resent repeated questions or probes. When this happens, their answers can become stilted and lead to an uncomfortable and unproductive experience. But as Collier and Collier (1986) argue, materials such as photographs can function as a third party in interviews, as researchers and informants work together to explore and understand their content. They suggested that “verbal questioning can create a distance between interviewer and informants, whereas use of photographs can pull people together” (p. 131).

Not only can such tasks make interviews more comfortable, they also can equalize the interpretive process by giving participants greater voice (Walker & Widel, 1985). Research relationships often are fraught with power imbalances, such as those between scholars and practitioners, adults and children, men and women, Whites and minorities, or members of different cultures or economic backgrounds (Schratz & Steiner-Löffler, 1998; Tobin & Davidson, 1990). Participants can be keenly aware of these imbalances, and they may actively resist researchers’ attempts to get to know them (Tobin, 2000). But by sharing in the definition of meaning, individuals no longer serve as the “subjects” of research but take on a more active role and exert greater agency (Harper, 1988). In some cases—such as the analysis of photographs—they literally become experts who lead researchers through content (Collier & Collier, 1986) rather than receptacle of information, extracted by researchers, analyzed and assembled elsewhere (Banks, 2001). With many tasks, participants can control the exchange of information, introduce ideas and concepts they consider relevant and significant, and participate meaningfully in the research process.

Giving participants greater control can also yield data that more authentically reflect their conceptual categories—for example, by revising, adding to, or omitting stimulus materials (e.g., Barton & McCully, 2005; Cunningham, 2006). Allowing for such input not only shifts the power balance but also makes data more valuable. Interpretations constructed by outsiders may have little explanatory power in naturalistic contexts. We can often understand real-world settings only by becoming familiar with how insiders make sense of their world. This, after all, was Jean Piaget’s central project: to understand the development of children’s thinking on its own terms—their “natural mental inclination”—rather than by forcing it into adult frameworks of interpretation (Ginsburg, 1997, p. 48). Although researchers may believe that they can
Elicitation Techniques

assemble tasks that allow perspectives to emerge, asking participants themselves to contribute to the process makes this more likely. Allowing participants to construct, name, and explain their own conceptual categories increases the likelihood that responses will be grounded in their own experiences and that they will supply ideas and associations that researchers might not notice on their own (Gates, 1976; Harper, 1988; Schratz & Steiner-Löffler, 1998; Soley & Smith, 2008; Wagner, 1979).

ARRANGEMENT TASKS

One of the most useful forms of elicitation involves asking participants to categorize or sequence stimuli and explain their arrangements. By making judgments about how items relate to each other, participants are led to consider the abstract principles, patterns, or conceptual categories that guide their decisions—precisely the kinds of tacit understandings that otherwise may be difficult to articulate. This kind of task is especially useful when researchers hope to avoid imposing their own concepts on participants—and thus overdetermining their responses—and instead want to give them the chance to articulate categories that make sense to them. This is part of the long-standing quest in anthropology and other social sciences to better understand how participants themselves see the world (Bernard, 2006; Borgatti, 1999; Weller & Romney, 1988).

Sorting

Sorting tasks—are also referred to as card sorts, pile sorts, and judged similarity tasks—are used to elicit participants’ conceptual categories. The purpose of sorting is to identify how participants think about which items go together in a category and why (Bernard, 2006; Borgatti, 1999; Roos, 1998; Weller & Romney, 1988). A sorting task begins with a set of items (often on pieces of heavy or laminated paper to encourage physical manipulation), such as words or images that refer to physical objects, historical or contemporary events, teaching practices, or any aspect of culture and society. Participants are asked to sort these into groups that make sense to them, so that the items in each pile belong together in some way. Usually participants are not given specific criteria to use in sorting items, although they may be asked to sort them into a given number of groups (because some people tend to be “lumpers” and others “splitters”). After settling on their groups, participants may then be asked to label each group and then, crucially, to explain how they made their decisions about which items should go together. A variation of sorting involves systematically presenting two items at a time and asking participants to explain which two are most similar and why.
Barton and McCully (2005, 2010), for example, used a sorting task to investigate how students in Northern Ireland connected history to their senses of identity. Previous research on young people’s identities in Northern Ireland had relied on written surveys in which students were simply asked to select from among terms, such as British, Irish, Northern Irish, or European. This kind of survey item is not explicitly linked to history, but just as importantly, it assumes that researchers can determine ahead of time the alternatives that students will consider most salient. To allow greater space for students to discuss their own concepts of history and identity, Barton and McCully provided participants with 28 images covering a variety of historical topics, which included British, Irish, Northern Irish, and world history; political, military, and social history; iconic Nationalist and Unionist leaders; as well as individuals unconnected to the two countries. They asked students to work in pairs to assemble these into “groups that you think belong together” (2010, p. 179). After students had completed the task, they were asked to explain what tied each group of images together, to select which image or group had “the most to do with you or who you are” (2010, p. 154), and to explain why. Many students created and selected groups that related to either Nationalist or Unionist history, yet the majority did not. Instead, they chose groups related to leaders, local history, or the Troubles more generally. Students’ attachment to such historical topics likely would not have been evident if researchers had determined the set of categories ahead of time.

Cunningham (2003, 2006) used a card sort exercise as part of her investigation of British history teachers’ instructional goals and practices. Drawing from long-standing controversies over the purpose of teaching history, Cunningham created 16 cards, each with a different sentence that captured important aspects of these perspectives. She placed each sentence on an index card and asked teachers to sort the cards into groups. She then asked them to identify those that came closest to their own goals for students. She found, however, that participants’ groupings had little connection to the controversies that so inspire historians, politicians, and teacher educators. Instead, they grouped them into such categories as primary and secondary goals, goals versus activities, goals and facilitating activities, and practices that are intrinsic to the subject versus those that enrich students’ lives. This is precisely the value of a sorting task: It can show that the concepts we develop as theorists may have little salience to practitioners, who see the world in very different terms.

Although anthropologists often use data derived directly from sorting tasks as ends in themselves, their usefulness is magnified when participants’ responses become a stimulus for conversation. Barton and McCully (2005, 2010), for example, used students’ groupings as an entry point for asking them about perceptions of history in Northern Ireland. The images and groupings were concrete reference points in these discussions—ones that students referred back to often—and the sorting task itself appeared to serve as a kind of “warm up” to more abstract discussions. For Cunningham (2003, 2006),
the sorting task was just one component of a study that included multiple interviews and long-term classroom observations. In some cases, teachers’ observed practices contradicted the goals they claimed to hold when sorting cards, and discussing such discrepancies led to important insights into factors that mediated teachers’ ability (or desire) to enact their purposes for instruction.

**Ordering**

Another useful strategy for exploring participants’ conceptual categories is ordering. As with sorting tasks, participants are given a set of items, but instead of placing them into categories, they are asked either to arrange them along a continuum or to select some subset as being most central to a concept, and then to justify their choices. Depending on the topic, participants might be asked to order items based on importance, interest, difficulty, age, or any other characteristic that varies along a scale. This is a familiar task, because most people have practice talking about how things have more or less of a certain quality and why they think so—the best recipes, the funniest movies, the most dominant athletes, and so on. Even young children are often asked such questions as “Which do you like best?” and asked to explain why. In informal settings, justifying this kind of ordering (even though no one calls it that) is the foundation of much conversation. Participants are familiar with this format and can easily understand what is being asked of them.

To investigate elementary students’ understanding of historical time, for example, Barton and Levstik (1996; see also Levstik & Barton, 1996) presented participants with two large photographs—one of a family moving west in the late 1800s, the other of teenagers at a drive-in restaurant in the 1950s—and asked them to sequence the images based on which was older and which was newer, or “closer to now.” Once students had decided on the order, they were asked to explain how they knew. Then they were presented with seven other images, one at a time, and for each, they were asked whether it was older or newer than the others, from about the same time, or between two of the times, and to explain their reasoning. This research yielded important information on students’ chronological understanding (even 6-year-olds put most images in the correct order), the content of their historical knowledge (most of their explanations related to details of social life, such as fashion or technology), their reasoning processes (historical changes were tied to a narrative of progress), and the structure of their understanding of time (sequencing preceded knowledge of specific dates).

Ordering need not always involve placement along a continuum. In Ho’s (2010) study of Singaporean students’ understandings of the country’s national narrative, she presented students with a set of captioned images that reflected a variety of people, events, or ideas from the country’s past and present and asked them to select those that would best represent Singapore to someone
from another country. She then asked them to explain their choices and omissions. She found that despite differences in students’ ethnic, academic, and socioeconomic backgrounds, they consistently represented Singapore’s history in terms of progress and consensus, and they particularly emphasized the need for harmony in the face of potential racial and religious conflict. Ho found few instances of countervailing perspectives or attempts to think outside the rigid racial categories that characterize official policy in Singapore. The uniformity of students’ responses suggests that despite Singapore schools’ espoused goal of developing “critical thinkers,” the curriculum had not provided them with a way of thinking about alternative historical narratives. Similar tasks have been used to investigate students’ ideas about historical significance in a number of countries (e.g., Barton, 2005; Barton & Levstik, 1998; Levstik, 2001; Peck, 2010; Yeager, Foster, & Greer, 2002) and among varied ethnic groups in the United States (e.g., An, 2009; Epstein, 2009; Terzian & Yeager, 2007). Such studies have provided important information on the socially situated nature of students’ perceptions of the meaning and purpose of learning about the past.

Ordering tasks are especially useful for determining which features people consider central to particular concepts or practices (such as historical significance or the ordering of events in time). Stated another way, rankings are useful in eliciting the underlying criteria people use to determine how well an item fits certain conceptual categories. Although the tasks in these examples investigated very different aspects of students’ understanding, each required students to rank items and then justify their rankings. In each, the data demonstrated criteria that students considered salient in making their decisions—details of social and material life in the case of historical time, progress and consensus in the case of national identity. Similar tasks could be used to investigate activities that respondents consider most central to citizenship, equality, cultural harmony, economic development, or other concepts. On a practical note, however, researchers should not expect respondents to explain each of their choices and their relative rankings when working with a large set of items. Providing such detail is too tedious, and it suggests a higher degree of precision than respondents are likely to have applied in their rankings.

CONSTRUCTION TASKS

Instead of asking participants to manipulate materials created by researchers, a second group of techniques engages them in constructing their own responses to stimuli—in words, images, or a combination of the two. This provides an even greater space for uncovering how respondents make sense of the world, because they are less limited by available materials. Concept maps are a well-known example of asking participants to construct responses using words and images (e.g., Edmondson, 2000), but other techniques include drawing, projective devices, and freelistening and sentence completion.
Elicitation Techniques

Drawing

Asking participants to draw a picture or diagram has two important uses. First, drawing can be used with populations whose oral and written language may not be adequate to fully express their ideas, such as young children. To explore Irish primary children’s understanding of identity, for example, Waldron and Pike (2006) asked them to draw pictures of what it means to be Irish—both from their own perspective and from that of an imagined person from another country. They also asked children to write and talk about this issue, but drawing provided an alternative means of expression and thus potentially tapped into the thinking of students who were less verbal or who were less comfortable with writing or being interviewed. Drawing is a highly familiar task for most children, and even the youngest are unlikely to balk at a request to “draw a picture of . . . .”

Drawing is especially useful in investigating ideas about geographic space, and it is an established technique in research on children’s geographic thinking (Pike, 2011). Gillespie (2010), for example, asked Amish and non-Amish children to draw maps of their neighborhoods. Amish children focused primarily on their families’ farms and excluded those of their neighbors, while non-Amish children (in the same area) tended to include neighbors, as well as a variety of local landmarks beyond their own farm. Other written or verbal methods would be unlikely to provide such insight into students’ socially situated constructions of geographic space.

Although drawings may seem appropriate only for young children, this method can stimulate deeper and more elaborate verbal responses even among older students and adults. Whenever people create art, they make symbolic choices, and interviewers can ask them to reflect on the meaning of those choices. Hunter and Farthing (2008), for example, asked teenagers to create representations of concepts such as nation, heritage, interpretation, memory, and identity. One student, who focused on culture, drew a fruit bowl and wrote:

Culture is like a fruit bowl, full of lots of different flavours, shapes and sizes, each unique in its own way . . . . It is all piled together in the world, changing with age, depleting in its uniqueness, as it becomes mouldy, more like each other. Many cultures are dying in their traditions and heritage today, like fruit as it rots. (p. 17)

This was clearly an articulate and creative student, but the depth of her answer may have been due in part to the need to justify the symbolic choices she made in her drawing. Similarly, Bagnoli (2009) asked young people who had migrated between England and Italy to draw who they were at that moment and to include people and things important to them. The visual elements of their drawings often served as metaphors that allowed them to holistically express
their associations and meanings and to narrate complex stories about their lives—as in the work of one youth who used his drawing of a crossroads to explain the choices he faced in life.

Like many elicitation techniques, drawings are valuable in large part because they lead to better conversations. Having children draw a picture, for example, can serve as an icebreaker that makes them feel more comfortable in interviews (Morrow, 1998). Older participants may also be put at ease by beginning with a drawing, although some may be put off by the prospect of a researcher watching them as they draw. This is less problematic in group settings, where after some initial reticence, participants usually shed their inhibitions as they recognize they are not alone in lacking artistic skills. They quickly realize that it is not the drawing that is important but what it allows them to express. This expression of meaning is precisely what researchers want, and by asking participants to explain and reflect on their drawings, they gain insight into their conceptual understandings, which may include their “plans, dreams, dilemmas, and emotions” (Bagnoli, 2009, p. 551).

**Projective Devices**

Projective devices include a large class of tasks and procedures that ask participants to respond to ambiguous or open-ended stimuli, sometimes by imagining themselves in a particular setting. In popular culture, the most well-known such instruments are the Rorschach test (in which participants describe what they see in a standardized set of inkblots) and the Thematic Apperception Test (in which participants tell dramatic stories about ambiguous images of people in daily settings). Psychologists have most often used such techniques to assess personality characteristics (or in some cases to identify a history of experiences such as child sexual abuse), but the validity and reliability of projective devices for clinical purposes has not been well supported in the literature (Lilienfeld, Wood, & Garb, 2000). Educational researchers, however, rarely use the standardized instruments and scoring indexes favored by psychologists, and they typically are not interested in personality types, unconscious psychological dynamics, or hidden meanings. Instead, they usually create their own content- and context-specific tasks to elicit, in a relatively straightforward way, attitudes about teaching, learning, or the wider society.

Recent theorizing about changes in youth political involvement, for example, suggests that young people prefer activities that involve social welfare needs rather than community building and that they are more likely to prefer volunteerism and direct political action over traditional forms of political involvement. Middaugh (2009) investigated the extent to which these variables influenced students’ willingness to engage in civic and political activities, as well as whether their attitudes varied by gender or ethnicity. Her research design was a traditionally quantitative one, but her instrumentation could not
Elicitation Techniques

involve simple attitude scales, because it was unlikely that her participants would be working with well-developed conceptual distinctions among these features of civic and political involvement—and even if they were, it was unlikely that the narrow and precise wording of a survey could capture those distinctions. Instead, she used a series of scenarios in which conditions of civic involvement were systematically varied on the basis of whether they involved direct action, volunteerism, or traditional political activity (such as collecting signatures for a ballot initiative). For each, she asked whether the character in the scenario should help with the planned action, whether it would be acceptable for the character not to participate, and whether the participant him/herself would take part. Somewhat surprisingly, she found that young people in her study did not prefer social welfare issues over community building, that they considered volunteerism and traditional politics equally desirable, and that they preferred both of these to direct action. The concreteness of these scenarios likely produced more specific responses than abstract questions about “social welfare” or “community involvement” could have, and asking students to evaluate the proper behavior for both themselves and others provided information on the extent to which they considered civic involvement obligatory.

A similar method can be used to investigate how students make sense of the perspectives of people in history. This is a key aspect of historical thinking: To understand the actions of people in the past, students have to understand their values, attitudes, and beliefs, as well as recognize how societal institutions at the time structured people’s opportunities and limitations. Students must also attempt to distinguish historical patterns of thought from those that characterize contemporary society. Historical perspective-taking has been the object of a great deal of scholarship in the field (for a review, see Barton & Levstik, 2004), but empirical research has been limited by a reliance on reactions to verbal questions and written sources. Although it is possible to tease out students’ understanding of perspective from their answers to questions about the actions and motivations of people in the past, this kind of analysis involves a high level of inference, precisely because this is a topic that students often have little experience discussing. Gambill (2014), on the other hand, presented students who were participating in a simulation of the Underground Railroad with images such as that in Figure 1, and she asked them to imagine what the character was thinking. Students are familiar with the idea of thought balloons, and this task allowed them to talk about empathy in a way that was both concrete and familiar. She found that students’ understanding of the factors involved in deciding to escape were more varied after the simulation than before and that high school students’ ideas were more complex than those of middle school students. Asking follow-up questions such as “Is that what she would think or what you would be thinking?” or “What about if you went back in time, what would you be thinking?,” would provide further data on the distinctions students made between historical perspectives and their own, and such questions would be easily comprehensible within the context of responding to the image.
Projective techniques can be particularly useful when researchers are interested in topics that are socially sensitive. Such techniques can allow participants to displace their thoughts or feelings onto others and to thus make their own disclosures seem potentially less threatening (Soley & Smith, 2008). These tasks have been shown to produce answers that are not only more elaborate but materially different than first-person questions (Soley & Smith, 2008), and they may capture attitudes or beliefs that participants otherwise would be reluctant to reveal (Catterall & Ibbotson, 2000). Two important but under-researched topics, for example, are ideas about religion and about economic inequality. Bringing up such topics can be controversial, and asking children about them could provoke anxiety, but projective techniques can alleviate some of this potential awkwardness. In a study of the development of students’ ideas about religion, for example, Goldman (1964) showed (Christian) students a drawing of a young child entering church and asked how often they thought the family went there, what they thought the child enjoyed and did not enjoy about church, why the mother and father attended, and whether going there helped the child in any way. He then showed a drawing of either a girl or boy praying beside his/her bed and asked children how often they thought the child prayed, what they thought the prayer was about, to whom it was addressed, whether
prayer requests ever came true, and why. Displacing the questions onto children other than the respondents themselves may have made this a more comfortable task than direct questions about their own religious ideas and experiences.

Similarly, to investigate young people’s views about socioeconomic status, Weinger (2000) showed participants from lower- and middle-class economic backgrounds photographs of two homes—one of them rundown and unkempt, the other in good condition and with a well-manicured lawn—and asked them to describe the people who lived there. She found that although both groups characterized the imagined people in the rundown home in positive or neutral terms, middle-class children’s responses were vague, while poorer students specifically identified characteristics such as family closeness, coping skills, and helpfulness, alongside financial hardships and social rejection. Both groups also described the middle-class family in positive terms, but middle-class children were more likely to point to their normality and responsibility, while poorer children also noted their financial good fortune, rule-following, cleanliness and good manners, and quality of parenting. Some, though, suggested they were callous and indifferent to the plight of the poor. This task likely was effective in eliciting participants’ responses not only because of its visual nature but also because it asked students to imagine personal characteristics of imaginary families. Students may have been more reticent to talk about their own families or those of people they knew.

Freelisting and Sentence Completion

One of the simplest forms of elicitation is freelisting, in which participants name (either orally or in writing) all the examples of something they can remember. Anthropologists, for example, have asked participants to list examples of animals, fish, fruit, illnesses, and even “bad words.” Because some cultural groups divide the world differently than others, this task provides information on what a community considers fruit, illness, bad words, and so on. Asking participants to list items provides information not only on boundaries of a cultural domain (what is a bad word and what is not) but on how important particular examples are within that domain. Researchers who use this technique assume that the more widely shared an example is (i.e., the more often it crops up on participants’ lists), the more central it is to that group’s way of thinking about the domain. In addition, the earlier an item occurs within lists, the more salient it is within the community (Bernard, 2006; Borgatti, 1999; Weller & Romney, 1988).

Seixas (1994), for example, gave 10th-graders a survey in which he asked them to do the following: “Think briefly about all of the events and developments which have happened in the past approximately 500 years. List three which, in your opinion, are the most important” (p. 301). Using this freelisting task, he was able to identify not only the specific events students considered
important (their top three choices were the World Wars, the rise and fall of Communism, and European exploration of the Americas) but also the kinds of events they thought were significant: Nearly all responses related to international politics and war, whereas very few dealt with gender, reproduction, work, childhood, diet, art, or thought. For these students, history (or at least, significant history) consisted of war and politics, not of social and material life.

Seixas also interviewed a subset of participants to explore their reasons for identifying items on their lists, and in these discussions, it became clear that students used two principal criteria for judging historical significance: Some items were part of narratives that explained the impact of events on the modern world, while others served as analogies from which people in the present might learn lessons.

Because of the simplicity of freelisting, it can be used with large samples and can produce extensive numerical data. Wineburg and Monte-Sano (2008), for example, asked 2,000 high school students and 2,000 adults to name the most famous Americans in history, from the time of Columbus to the present. (They asked participants not to include presidents or their wives.) Interestingly, each of the top three responses among students were African Americans (Martin Luther King Jr., Rosa Parks, and Harriet Tubman), and each of these were among the 10 most frequently named figures among adults. This suggests that the inclusion of African American history in both schools and popular culture has achieved a remarkable degree of consensus in the contemporary United States. At the same time, other struggles for equality or reform went largely unnoticed by both groups: Susan B. Anthony was the only suffrage leader to be mentioned on more than 5% of the lists, and other reformers, labor activities, or minority leaders were named by only a handful of students. These findings led the authors to suggest that “America’s multihued movement for equality . . . has been seemingly reduced to an equation of black and white” (p. 646).

Another study of the same topic illustrates some of the trade-offs involved in designing such tasks. Epstein (2009) asked 46 high school students to both identify and explain three important events and actors in U.S. history. Analyzing the amount of data produced by such explanations required a much smaller sample size and thus allowed for many fewer comparisons, but the result was a more nuanced understanding of differences among groups. Wineburg and Monte-Sano (2008) found differences in the frequency with which Whites and Blacks identified members of each ethnic group, but Epstein was able to describe the nature of their thinking: Whites were more likely to see civil rights as bestowed by Whites upon Blacks, whereas Blacks thought of the struggle as having been waged by members of that community and as having been longer in coming.

Sentence completion shares many similarities with freelisting. Rather than asking for a list of items, however, researchers provide a stem and ask participants to complete the sentence with single words and phrases (“I like to learn
about . . . ”) or more open-ended and potentially elaborate responses (“School is . . . ”). Longer stems typically are more structured and narrow the range of possible responses, while shorter stems provide more freedom for participants to interpret sentences in a variety of ways (Soley & Smith, 2008). Although this technique is sometimes classified as a projective device and used to measure personality characteristics or emotional disturbances (Forer, 1960; Rotter, 1951), it has also been used to explore participants’ social attitudes and their ideas about specific educational topics. In evaluating the implementation of a new curriculum in New Zealand, for example, Sinnema (2010) asked teachers to complete the sentence, “Values in the New Zealand Curriculum requires teachers to . . . .” The intent of the curriculum was for teachers to provide students with opportunities to explore the nature of values and their role in social life, but Sinnema found that many respondents focused only on instilling particular values, especially as a means of behavior management.

Sentence completion tasks can also be used when researchers want to narrow and focus participants’ responses. In research on students’ ideas about historical agency, for example, Barton (2010) found that when advanced history students were asked (without an elicitation task) to explain the rise of Nazism, many of them provided detailed and sophisticated explanations of a variety of historical contingencies. They pointed to such factors as ideology, popular feeling, political alliances, economic fluctuations, the intransigence of political leaders, and the weakness of international agreements. Their explanations were so sophisticated they left almost no stone unturned. But when asked instead to complete the sentence, “The rise of Nazism would never have happened if . . . .” their responses became much more pointed and specific, and this provided insight into what students considered necessary and sufficient causes rather than supporting and catalyzing conditions. A number of students also implied that nothing could have prevented Nazism, and this suggests a view of historical inevitability that ignores the ability of groups and individuals to influence the course of history—a finding that has important implications for how students interpret the role of agency in social life.

Freelisting and sentence completion share a number of important practical advantages: they can be adapted to practically any topic; are easy to administer; can be used in both oral and written versions; can be used with groups or individuals; can be adapted to a variety of cultural contexts; and produce data that can be analyzed qualitatively, quantitatively, or both (Lindzey, 1961; Soley & Smith, 2008). Regardless of format, though, these techniques are best used when the topic of interest has an accepted name or can be clearly described (Borgatti, 1999), such as “famous Americans,” “important historical events,” and “what a citizen does.” However, freelisting can be alienating when respondents feel that researchers are testing their knowledge rather than exploring their ideas, and thus, topics with a limited number of correct answers should be avoided. Asking participants to list World War II battles, or to complete the sentence “Most immigrants in the 19th century came from . . . ” will make
them feel as though they’re being quizzed, and they will shut down rather than open up.

In addition, even when explanations are asked for, the terse nature of these tasks does not usually lead to in-depth data on the knowledge, attitudes, or perspectives of individuals. Freelisters in particular are better used to create an aggregate portrait of a population, as in the studies described here. Because they provide group-level data, free lists are also particularly suited for comparisons of populations, and this is an important feature of studies by Epstein (2009) and Wineburg and Monte-Sano (2008). Researchers investigating teachers’ ideas about civic participation, then, probably would not use freelisters to examine individuals’ ideas, but they could ask all teachers at a school (or throughout a district) to list “all the things you think students need to know about democracy” or “all the behaviors that students might display to indicate that they can act politically.” These would lead to a portrait of shared assumptions and could be used to compare settings—by collecting data in urban and suburban schools in the same area, for example, or among educators in different countries or regions. As Bernard (2006) observed: “The humble free list has many uses. Use it a lot” (p. 305).

EXPLANATION TASKS

In the final group of elicitation techniques, participants are asked to explain the content of materials (usually visual images) or to describe a process (such as their own thinking or teaching). In these tasks, stimuli are used primarily to prompt thoughts that are largely tacit or rarely articulated. Researchers frequently are interested in precisely these kinds of ideas—concepts and mental procedures that are hard to understand or access without participants’ explanations, but which they take so much for granted that they find them hard to explain. Three of the most well-known such tasks are thinking aloud, stimulated recall, and photo-elicitation.

Thinking Aloud

Two of the more familiar forms of elicitation techniques involve “introspective methods” in which participants reflect on the nature of their thinking. The first of these methods involves what are variously known as “think alouds,” “verbal reports,” or “protocol analysis.” In this approach, participants are asked to report on their thinking, either while they are in the process of conducting some task or immediately afterward. This approach is particularly useful when researchers hope to gain insight into how participants engage in some cognitive process, such as problem solving, decision making, second language learning, or text comprehension (Ericsson & Simon, 1993; Pressley & Afflerbach,
In an individual meeting with the researcher, participants are asked to think aloud as they work on a problem (or read some portion of text) and to say everything that comes to mind while engaging in the task. The goal is not for them to analyze or explain their thinking but simply to state aloud the thoughts that occur to them. Participants are asked to speak constantly, and the researcher’s intervention is usually limited to reminding them to speak.

Although most think-aloud research involves adults, children are also capable of verbalizing their thoughts while engaged in reading or other tasks. To investigate the effect of classroom investigations with historical evidence, for example, VanSledright (2002) asked fifth-graders to think aloud as they examined sets of written and visual sources. Before instruction, students’ verbal reports indicated that they focused primarily on comprehension monitoring strategies (such as rereading or summarizing) and intratextual analysis (reflecting on the content of the source, the author’s style, and their own reactions). After instruction, students were more likely to engage in intertextual analysis, such as comparing evidence, assessing the perspective of authors relative to each other, and drawing conclusions based on evidence. When researchers are interested in how students mentally work through a task, such as analyzing primary sources, think alouds are an especially useful technique.

Think-aloud research is generally done in a single sitting, with relatively brief tasks or texts, but it can also be used in more complex ways. VanSledright and Afflerbach (2000), for example, were interested in preservice teachers’ cognitive strategies and affective responses when reading revisionist history texts. These texts were too long and demanding for a typical face-to-face interview session, so the researchers met with participants to talk about the process of thinking aloud and asked them to conduct recordings on their own time over the course of the next month. They were asked to talk about how they were trying to construct meaning, including how they reconciled texts with each other and with their prior knowledge, as well as the sorts of reading, reasoning, and thinking strategies they used as they did so. The researchers found that participants modified their understanding of the historical period and gradually assimilated and partially appropriated authors’ narrative accounts. In doing so, they monitored their understanding of text, considered prior knowledge, and evaluated authors’ claims. Methods that only compared teachers’ thinking before and after reading might have detected changes in their ideas, but they would have been less able to identify the processes by which those ideas changed. Examining such processes is an important, and largely unexplored, avenue for research on teaching and learning in social studies (Barton, 2008).

Stimulated Recall

A second introspective technique involves stimulating participants’ description of mental processes after the fact. In many naturalistic settings,
participants cannot be asked to verbalize their thinking at the same time that they are involved in ongoing activities. Teachers and students, for example, cannot easily think aloud at the same time they are engaged in lessons. As a result, researchers have attempted to gain insight into cognitive processes by asking participants to recall their thinking at a later time, frequently while they watch video recordings of lessons or lesson segments (Ericsson & Simon, 1993; Lyle, 2003; Yinger, 1986). As Yinger (1986) notes, although the ability of such recall to accurately mirror thoughts that had occurred during recorded events has been widely criticized, such procedures nonetheless provide important data on participants’ reflective understanding of the nature of teaching and learning. Stimulated recall cannot necessarily capture teachers’ thinking while in the process of teaching, that is, but it can nonetheless spur reflection on teaching by providing access to how teachers make sense of teaching episodes. By prompting explanation and justification of practices, it elicits teachers’ implicit theories and otherwise tacit beliefs.

Although stimulated recall has been used primarily in studies with teachers, it began as a method for conducting research with students, and it continues to hold potential for providing insight into the process of learning. Nuthall (1999), for example, asked upper elementary students who had been studying science and social studies to recall their experiences and thoughts while they watched video clips of their involvement in classroom activities. They were asked such questions as “What was going on for you here?,” “What’s this like for you?,” and “What’s happening here?,” along with probes that encouraged them to talk more about their experiences and thinking (p. 13). He found that students talked about activities primarily in terms of getting tasks completed (usually looking for the shortest and most efficient way to complete requirements) and seemed unaware of their intended academic purposes. They even viewed their teachers’ involvement as little more than annoying interference. In addition, students did not clearly differentiate their thinking from physical completion of the task (or from accompanying discussions with other students). And once they had completed the activity, their thinking stopped. They may have been surprised or confused by the solution they worked out, but resolving that confusion was irrelevant to completing the task, and thus new information was not fully integrated with their previous beliefs. Nuthall concluded that in the context of classroom activities, “students act and perceive rather than reflect and interpret” (p. 38). He further maintained, however, that students’ thoughts during the activities (including their confusions) were still available to them for later reflection and that it was often at these later moments rather than during the original experience that learning actually occurred—including during the stimulated recall interviews themselves. This is an important reminder that interview techniques themselves can hold important intellectual benefits for participants, particularly when they are students who are learning about the concepts being investigated.
Photo-Elicitation

A simple but productive technique involves showing photographs to participants (or asking them to select their own) and then asking them to talk about what they see. Anthropologists often take photographs of local people, places, and events and ask members of the community to talk about them. Their reactions can reveal community values, attitudes, and beliefs, as well as the meaning that participants attribute to aspects of the local setting (Collier & Collier, 1986; Harper, 2002; Wagner, 1979; Wang & Burris, 1997). Photographs can also trigger memories, and participants’ responses can illustrate group dynamics or provide insight into how social and cultural systems operate (Prosser & Schwartz, 1998). In schools, photo-elicitation can serve as a more concrete (and potentially more engaging) alternative to “grand tour” questions in which participants are asked to describe a typical day in the setting.

This technique can be particularly useful for surfacing taken-for-granted ideas about the social world and for revealing features of community life that researchers might otherwise overlook (Carlsson, 2001; Harper, 1988; Schwartz, 1989). In one study, for example, a teacher who was explaining a photograph of herself working with students noted a girl touching the teacher’s bracelet, and this led her to reflect on students’ interest in getting to know their teachers personally. This reflection might not have occurred without the presence of a small detail invested with great significance (Walker & Widel, 1985). Collier and Collier (1986) refer to this as the “can-opener effect” and note how photographs can speed up the process of establishing rapport. Photographs can release anecdotes and recollections that might otherwise remain buried. They further note that because photographs often are emotionally charged, participants sometimes use them as a stimulus for dialogues with themselves, and this leads to the kind of fluency and deep reflection that might not come about when responding only to verbal prompts. Collier and Collier even suggest that it can be harder for participants to lie about their reactions to a photograph because of their emotional salience.

Asking a variety of participants to react to photographs can be particularly useful in eliciting the range of ideas that exist in a setting, because not all participants will respond in the same ways (Banks, 2001). Their reactions and explanations, then, can provide information on how age, role, status, or other social and individual factors influence perspectives and how similar events are interpreted differently within and across communities (Schwartz, 1989). Prosser (1992), for example, showed British teachers photographs he had taken in a variety of everyday settings at their school. One was a photograph of a language arts text, *Nine Modern Poets: An Anthology* (Black, 1966), which had been vandalized by a student so that its title appeared as “Nine Nude Puffs in an Orgy” (*puff*, also spelled *poof*, is derogatory British slang for male homosexuals). The head of the school, an “‘elderly statesmen’ type” (p. 403), found the
image threatening, and he thought that showing the book (or the photograph) to students would lead to further vandalism. An art teacher, on the other hand, considered it an act of creativity and wished that more students demonstrated such talent. Notably, none of the teachers—or the researcher—commented on what the image had to say about gender and sexuality or whether it might reflect a climate of homophobia in the school or the wider society. As in all research, what participants fail to say can be just as enlightening as what they do say.

Photographs for elicitation can come from a variety of sources and be used for differing purposes, although researchers usually select images that they assume will have some meaning or significance for participants. These may be generic photographs of sites within a community (such as classrooms or corridors within participants’ own school), or they may include more specific images of participants themselves (as in the above example of the teacher working with her students). Photographs may also be chosen to be deliberately provocative or disruptive to elicit otherwise suppressed views, as in the example of the vandalized book. Images, however, do not have to be contemporary. In one study, teachers who examined photographs of themselves in earlier stages of their careers commented on how their appearances were linked to their professional identities as teachers at the time (Mitchell & Weber, 1998).

This kind of visual elicitation can also use moving images, as in the “video-cued multivocal ethnography” of Tobin et al. (2009). They recorded an entire day of activity in preschools in China, Japan, and the United States and edited each video down to 20 minutes. They specifically looked for scenes that had the potential to stimulate discussion and highlight disagreements about the mission of preschools and the nature of children, such as scenes of conflict, tension, separation, and intimacy (Tobin & Davidson, 1990). They first showed these recordings to staff at each of the schools and asked them to explain what was going on, assess whether it was typical, and evaluate their response to each situation. They then showed the videos to preschool teachers at other sites in the same countries and to staff of preschools in the other countries, and they asked for their reactions and for ideas about how they would have handled the situations themselves. Participants often had different understandings not only about what teachers should do but also about what was going on in the videotape. In their comments, participants illustrated not only their individual ideas but also the culturally and historically embedded constructions of childhood and schooling from which they drew.

Photo-elicitation is an open-ended and transparent technique, but whenever researchers select stimulus materials, it limits participants’ ability to represent their experiences (cf., Eldén, 2013). A form of photo-elicitation that has become popular in recent years aims to provide greater participant input into the research process. This approach, usually known as photovoice (Wang & Burris, 1997), asks participants themselves to take photographs that have meaning for them and then to explain their selections. This method is often associated with participatory action research in which participants are asked to
identify important needs in their community and to develop ways of addressing them. Burke, Green, and McKenna (2014), for example, investigated students’ understanding of public spaces in their neighborhoods by asking them to take photographs of objects and places they considered “good” and those that “needed improvement” (p. 13). Students were then asked to caption the images and explain the meaning behind them. Tupper, Carson, Johnson, and Mangat (2008), meanwhile, used students’ explanations of photographs they had taken in their schools to examine the messages about citizenship conveyed there, as well as how students negotiated the use of these spaces for interaction and identity. In both these studies, starting with students’ own representations provided insights that researchers might not have been able to identify without insiders’ creation of materials.

**CONCLUSIONS**

Elicitation techniques provide alternatives to direct, verbal interview questions and are especially useful when researchers want respondents to talk about controversial topics or ideas they have little experience discussing. The methods covered here facilitate such discussions, particularly by using concrete and familiar tasks to explore abstract concepts, by shifting attention onto external materials or scenarios, and by reducing power inequities. These techniques are not foolproof, and their success is inseparable from other characteristics of good interviews, such as researchers’ rapport with participants and their ability to encourage conversation. Yet, with judicious selection, careful planning, sensitive implementation, and thoughtful consideration of the suitability of tasks for given topics, populations, and circumstances, elicitation techniques can help researchers create valuable data on topics that might otherwise be difficult for participants to discuss.

It is important to keep in mind, though, that such methods do not entirely transform the power relations involved in research. Even though participants may have a greater voice in some kinds of tasks than others, their experiences may still leave them feeling powerless about the extent to which others have analyzed their thoughts (Tobin & Davidson, 1990). Addressing such perceptions may require engaging participants not only in interesting and open-ended tasks but also in the design of research itself (e.g., Waldron, 2006). Ultimately, researchers must remember that their data derive from particular tasks, which are interpreted in particular ways, and that these are never free from the influence of power-laden social contexts.
REFERENCES


Gambill, R. (2014, November). “It’s not like they can beat you”: Living history experiences and historical understanding. Paper presented at the annual meeting of the College and University Faculty Assembly, Boston, MA.
Barton


Elicitation Techniques


**ABOUT THE AUTHOR**

KEITH C. BARTON is a Professor in the Department of Curriculum and Instruction at Indiana University, Bloomington, IN 47405. He can be contacted at kcbarton@indiana.edu.