The Effect of Language of Survey Administration on the Response Formation Process
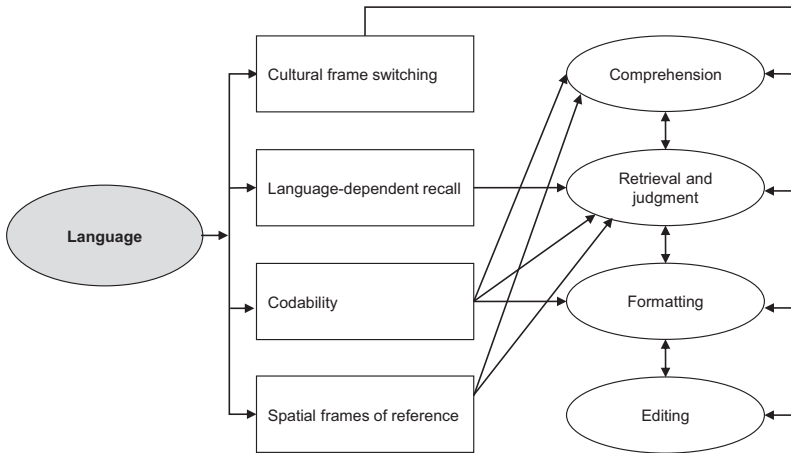
Emilia Peytcheva

Introduction

With the increasing number of people of multiple cultural backgrounds in modern societies, surveys of ethnic minorities and immigrants are becoming more common. One obvious source of measurement differences is the necessary use of different languages when intending to measure the same phenomena in multiple ethnocultural groups. Typically, surveys allow respondents to answer in the language of their choice, possibly introducing self-selection bias to the extent to which those who choose their mother tongue differ in background characteristics (e.g., level of acculturation, education), substantive answers, and response patterns (e.g., "don't know" responses) from those who choose the mainstream language. However, although self-selection certainly plays a role in differences observed across the different language versions of a survey, it is premature to consider it the sole source of all observed differences.

There is a known link between language and cognition (e.g., Whorf, 1956). To study language influences on the response formation process in surveys, we need to assert that the various language versions of a survey are free of translation problems and convey the same constructs. Thus, any observed differences between responses provided by the same respondent in different languages can be attributed to language priming a particular mind frame and influencing the thought processes.

To examine the potential effects of language on survey responses, we focus on the response formation model (Sudman, Bradburn, & Schwarz, 1996; Tourangeau, Rips, & Rasinski, 2000). The right-hand side of Figure 1-1 presents the tasks that respondents perform to answer a survey question: attending to the question and response options (comprehension), retrieving the necessary information (retrieval and judgment), assessing the

Figure 1-1. Effect of language on the survey response formation process

completeness and relevance of the memories (formatting), and editing the response before mapping to the provided response categories (editing). These tasks are not necessarily sequential or independent but are presented as such for simplicity. The left-hand side of Figure 1-1 represents the mechanisms related to language influences that are most likely to be present at each stage of the response formation process. We limit our discussion only to mechanisms well known to yield reporting differences, namely, cultural frame switching, language-dependent recall, language codability, and spatial frames of reference inherent in each language. We acknowledge that other language influences might be at play, but as of now, they remain undiscovered.

The language influences presented in the model can only be apparent among bilingual respondents, so we discuss them within the context of more than one language being available to communicate with respondents. We describe each of these mechanisms and examine their possible effects at each step of the response formation model by reviewing the existing literature from relevant fields and deriving conclusions about consequences for surveys.

Comprehension

Survey data are meaningless if respondents do not understand the survey questions as intended by the researchers. Question comprehension involves processing the syntactic structure and understanding the semantic (literal)

and pragmatic (intended) meaning. In cross-cultural surveys, in addition to the direct impact of translation, comprehension problems may occur as a result of differences related to cognition. Because language is a tool for information exchange among people of the same culture, it reflects the meaning system of the culture. Thus, word meaning and sentence meaning in language comprehension depend on preexisting background knowledge about not only the grammatical norms associated with the language, but also the cultural norms and practices related to it. Furthermore, lexical ambiguity is inherent in languages, and recall of the lexical meaning of words is often context dependent. Languages differ in their contextual dependency, and this difference is reflected in the conversational norms across cultures. For example, many words in Chinese acquire meaning only in the conversational context and cannot be translated word for word; this is related to the practice of East Asian cultures to read between the lines (for an overview, see Nisbett [2003]). Thus, the same question presented in Chinese or English to a bilingual respondent may convey a different meaning depending on how much contextual information is incorporated from previous questions.

Cultural Frame Switching

Differential context dependency can have consequences for question interpretation when partially redundant information is presented (e.g., Haberstroh, Oyserman, Schwarz, & Kühnen , 2002). In bilingual respondents, such context sensitivity is likely to depend on which cultural frame is primed by the survey question. Research on acculturation has demonstrated that individuals can possess more than one cultural identity (e.g., Berry & Sam, 1996; Hong, Morris, Chiu, & Benet-Martinez, 2000) and move between different cultural meaning systems, depending on situational cues and requirements. This phenomenon, known as "cultural frame switching" (Briley, Morris, & Simonson, 2005; Hong et al., 2000), is likely to have a strong effect on survey responding because each cultural meaning system serves as an interpretive frame that affects an individual's cognition, emotion, and behavior (Geertz, 1993; Hong, Chiu, & Kung, 1997; Kashima, 2000; Mendoza-Denton, Shoda, Ayduk, & Mischel, 2000).

Language can serve as a situational cue for the cultural system associated with it; thus, it may prompt bilingual respondents to differential question interpretation based on the cultural frame induced by it. Indeed, studies that have experimentally manipulated language assignment among bilinguals report responses consistent with the cultural system associated with the

assigned language (e.g., Peytcheva, 2019; Ross, Xun, & Wilson, 2002; Trafimow, Silverman, Fan, & Law, 1997). Such studies provide evidence that language is a powerful cue for the interpretive frame bilingual respondents adopt when answering survey questions.

Codability

Language codability is the ease with which a concept can be expressed in a language. Not surprisingly, the most highly codable concepts are presented by the most frequently used words, which are short and easy to write and pronounce (see Whitney, 1998). Codability affects cognitive processes such as retrieval (Lucy, 1992; Lucy & Shweder, 1979; Lucy & Wertsch, 1987) and comparative judgment (Kay & Kempton, 1984). However, codability may also influence question comprehension in surveys; the question target may be very different depending on whether a specific word exists in the language for a given attitude or behavior or whether several less specific terms are used to describe it. For example, in Chinese, there are separate terms for family members that have only one English equivalent—different words describe whether your "uncle" is your mother's brother or father's brother and whether he is a younger or older brother. Thus, it can be hypothesized that when asked in Chinese about two or more related people who can be labeled differently, respondents may think of them differently relative to when questions are asked in English when a common label is used. This difference may lead to inclusion errors when respondents are asked in English because of the failure to draw a lexical distinction across referents. Such interpretational differences across two languages may affect various respondent tasks in surveys (for example, household roster construction).

Spatial Frames of Reference

Languages have inherent frames of reference for describing relationships among objects. Psycholinguists distinguish between relative and absolute languages (also known as egocentric and allocentric). Relative languages, such as most Western languages, use a viewer-centered perspective, giving rise to descriptions such as "in front of me" and "to the left." Absolute languages use external reference frames, such as cardinal directions or an up–down axis; for example, speakers of Arrernte (Australia) will say "the fork is to the north of the spoon" (Majid, Bowerman, Kita, Haun, & Levinson, 2004).

Such intrinsic language differences may potentially affect comprehension in bilingual speakers of languages with different dominant spatial frames of

reference because these reference frames have been found to determine many aspects of cognition (see Levinson [2003]). Experiments by Pederson et al. (1998) demonstrate that the domination of a linguistic frame of reference in a language reliably correlates with the way its users conceptualize in nonlinguistic domains. For example, speakers of Mopan (Mayan) and Kilivila (Austronesian) cannot distinguish between two photographs of a man facing a tree when the position of the man and the tree are left–right mirror images of one another because such a relationship between the objects in both photographs is described as "tree at man's chest."

For survey practitioners, such findings suggest that speakers of languages that use different frames of reference may interpret survey visual images and response scales differently. For example, the orientation of a scale (vertical or horizontal) may influence how similar or distinct response categories are perceived, depending on the language used and its inherent frame of reference. However, such effects are likely to occur only in cases where the dominant frames of reference used in two languages are not functional equivalents of one another (as in the example with Mopan speakers where there were no functional equivalents of "left" and "right" in the described mirror-image photographs); thus, their impact on the survey response processes may be very limited. However, the relationship between dominant frames of reference and cultural orientation (individualistic vs. collectivistic; for reviews on documented social and cognitive differences, see Oyserman, Coon, and Kemmelmeier (2002) and Oyserman and Lee (2007)) remains unknown. To the extent to which ego-centered frames of reference are related to individualistic identities across cultures that use such languages and vice versa, the language of administration will be an important factor influencing survey responses. Similar to cultural frame switching, a speaker of languages that use different frames of reference would endorse more individualistic or collectivistic responses depending on the cultural identity evoked by the egocentric or allocentric frame of reference inherent to the language of survey administration. Such possibility deserves further investigation.

Retrieval and Judgment in Behavioral Reports

The information requested in a survey question is rarely readily available, and often respondents need to retrieve memories and assess their relevance on the spot. Because this process is somewhat different for behaviors and attitudes, we discuss each separately, starting with behavioral reports.

Behavioral questions often ask about past events that took place in a respondent's life. When such events have low frequency of occurrence or are of particular importance to the respondent, they may be directly accessible in memory (for reviews of issues related to asking behavioral questions, see Bradburn, Rips, and Shevell, 1987; Schwarz, 1990; and Strube, 1987). However, respondents often need to recall relevant information and count instances of occurrence (enumeration) or compute a judgment (rate-based estimation). The success of retrieving the information and its accuracy depend on time on task (e.g., Williams & Hollan, 1981), the elapsed time since event (Cannell, Miller, & Oksenberg, 1981; Loftus, Smith, Klinger, & Fiedler, 1992; Means, Nigam, Zarrow, Loftus, & Donaldson, 1989; Smith & Jobe, 1994), the availability and adequacy of retrieval cues (for a review, see Strube, 1987), and the match between the encoding and recall contexts (Tulving & Thompson, 1973). The context may vary from physical context (Godden & Baddeley, 1975; Smith, 1988) to mental and emotional states (Bower, 1981; Bower, Monteiro, & Gilligan, 1978; Eich, Weingartner, Stillman, & Gillin, 1975). Several studies have demonstrated that the language in which mental activity is carried out during information encoding creates an internal context analogous to a mental state and can serve as a retrieval cue during information recall; similarly, the language spoken aloud during an event creates an external context analogous to a physical context and can serve as a situational cue during event recall (Marian & Neisser, 2000; Schrauf & Rubin, 1998, 2000). Thus, a match between language of encoding and language of recall in surveys should yield more accurate responses among bilingual respondents.

Language-Dependent Recall

Language-dependent recall is the notion that the language may influence retrospective reports. This phenomenon has been demonstrated in several bilingual groups in terms of number of recalled memories (e.g., Bugelski, 1977) and time in life when the recalled events took place (e.g., Schrauf & Rubin, 1998). Going beyond earlier findings of language-congruity effects, Marian and Neisser (2000) investigated whether a match between language of encoding and recall facilitated retrieval because the language matched words used during the original event or because the language at the time of recall induced a more general mindset, resembling the processes assumed to underlie state-dependent memory. The results showed that the effect of ambient language was significantly stronger than the effect of word-prompt language, further "strengthening the analogy

between language-dependent recall and other forms of context dependency" (Marian & Neisser, 2000, p. 366).

The implication of such findings for surveys that involve immigrant and ethnic minority populations is that the choice of language of survey administration affects both the quality and quantity of recall. Specifically, first-language cues tap into first-culture memories, while second-language cues likely activate more recent memories. This suggests that language of survey administration in bilingual respondents may be switched throughout the survey, depending on life periods for which researchers are interested in collecting data. Additionally, bilingual immigrants or ethnic minorities are likely to use different languages in different life domains, for example, at work and at home. We can expect that the match between language spoken at home and language of survey administration will yield the most accurate information regarding home events, the highest number of such reported events, and the lowest response latencies for home-related questions, and vice versa. Such hypotheses, if supported, would further argue for a language switch across domains in surveys of bilinguals.

Codability

Often, there is no direct correspondence across languages with respect to terms that describe the same phenomenon; thus, using phrases or multiple words to describe the concept of interest is necessary during translation. Research related to language codability would predict difficulty in recall with difficult-to-code words because easily coded words (and, therefore, events associated with them) are remembered more easily (Lucy, 1992; Lucy & Wertsch, 1987). However, analogous to question decomposition, multiple words may provide more contextual cues that can ease recall and eventually improve report accuracy. To date, it remains unknown how such processes operate for users of two languages with different levels of specificity for the same concept.

Spatial Frames of Reference

A different aspect of language-dependent recall is demonstrated in studies of spatial cognition; the frames of reference used in a language to describe specific situations are likely to induce the same frame of reference in the nonlinguistic coding of the same situations (Levinson, 2003). Various experiments (Levinson, 2003; Pederson et al., 1998; Wassmann & Dasen, 1998) have shown that when speakers of languages with different dominant frames of reference are given

various memory and spatial reasoning tasks, the nonlinguistic frames of reference used to carry out these tasks match the dominant frames of reference of the languages (see Levinson, 2003; Pederson et al., 1998; Wassmann & Dasen, 1998). Specifically, speakers of languages that use absolute frames of reference (e.g., Balinese, Indonesia; Belhare, Nepal; Arrernte, Australia) preserved the absolute coordinates of objects when performing tasks such as memorizing order and direction of objects within an array, while speakers of relative languages, such as Dutch, Japanese, and Yukatek (Mexico), preserved the relative coordinates of objects (Levinson, 1996; Pederson et al., 1998).

The cognitive consequences of being bilingual in languages that use different frames of reference remain unclear. One possibility is differential perceptual tuning due to the use of different frames of reference because languages have been found to affect perception such that individuals become more or less attuned to certain features of the environment (Goldstone, 1998; Sloutsky, 2003). For survey practitioners, this may mean that what is reported during recall tasks may be related to what language is used during initial information encoding and later, during the survey interview. In an extreme example, certain information may not be encoded *because* of the language spoken during an event that predetermines on what speakers focus their attention. Furthermore, similar to language-dependent recall, it can be expected that a match between language frames of reference during encoding and retrieval could facilitate remembering.

Retrieval and Judgment in Attitudes

Attitude questions often require respondents to form an opinion on the spot in the specific context of a survey (Sudman et al., 1996). To do so, they need to form a mental representation of the question target based on the most accessible relevant information. Preceding questions, visual aids, and interviewer characteristics can make certain information more accessible; language of survey administration can also determine what information is accessible at any given time by activating the cognitive–affective cultural framework associated with it. By using a particular language, a "language-specific self" is activated, who acts like a filter through which information is both encoded and retrieved (Schrauf, 2000).

Cultural Frame Switching

Language can affect what information is temporarily accessible by evoking a particular mindset related to the cultural meaning system associated with it. For example, a study of Greek students attending an American school in

Greece showed that the correlation between the same attitudinal questions administered in English and in Greek was low for domains in which the Greek and American norms differed in what was considered socially desirable and high for domains in which the cultural values converged (Triandis, Davis, Vassiliou, & Nassiakou, 1965). Similar results were reported for English–Spanish bilinguals by Marín, Triandis, Kashima, and Betancourt (1983).

Another aspect of cultural frame switching relates to differences in how Westerners and East Asians organize the world: Westerners show preference for grouping objects based on taxonomy or common category membership, while East Asians prefer groupings based on relationships (Chiu, 1972; Ji, Schwarz, & Nisbett, 2000). Such grouping preferences can be manipulated by the language used during the cognitive task; for example, Ji, Zhang and Nisbett (2004) found that relationship-based grouping shifted to categorical when Chinese speakers from Mainland China and Taiwan were asked questions in English. Recent studies in psycholinguistics have also demonstrated that language can affect comparisons (Bowerman & Choi, 2003; Gentner, 2003), and to the extent to which languages classify according to different criteria, the extracted similarities also differ (Boroditsky, 2001; Boroditsky, Schmidt, & Phillips, 2003; Lucy & Gaskins, 2001).

These findings have several implications for surveys of bilingual respondents. First, the information that is accessible to form an opinion will vary depending on the language of survey administration. Hence, to achieve maximum equivalence of different language versions, open-ended questions should be avoided. Second, the same question can be perceived to have different affective characteristics depending on the language and cultural norms it activates; thus, more or less socially desirable opinions will be expressed, depending on language. Knowing in advance how cultures differ in terms of a question's affective characteristics may better inform questionnaire design, and various techniques can be used to reduce social desirability or sensitivity across language versions. Third, judgments can be language dependent because comparisons are based on culture-approved practices and how language systems are organized. Such hypotheses necessitate systematic investigation of language effects and the underlying dynamics across question types.

Codability

Studies in psycholinguistics have demonstrated that codability affects judgment. Kay and Kempton (1984), for example, showed that color-naming practices affect judgments of colors: speakers of Tarahumara (a Mexican

Indian language that does not have separate words for blue and green) differentiated among color chips on the blue–green color continuum based on their physical characteristics—namely, wavelength of reflected light. In contrast, English speakers differentiated among the same color chips based on labels, such as "shade of green" and "shade of blue." Thus, English speakers evaluated colors in terms of categories in which they were easily coded, while Tarahumara speakers, lacking such codability of colors, based their evaluations on physical characteristics. Similarly, Hoffman, Lau, and Johnson (1986) examined the extent to which the codability of personality description (existence of stereotypes) in a language influenced the impression about a person. The study found that terms that were readily available in the language led to stereotyped impressions, and participants were more likely to elaborate on the described person's characteristics using terms consistent with the stereotype than when a verbal label was not available.

Such findings may have implications for the use of scales in surveys of bilingual respondents. For example, scales may be judged differently depending on whether scale labels are easily codable in both languages. If label equivalents are not easily codable in one language, respondents may be more likely to consider solely the numeric values of the scale when making judgments, resulting in response differences across language versions.

Response Formatting

The ability to differentiate among response options may be influenced by language codability, and the stimuli used to anchor the points of a rating scale may be affected by the cultural meaning system primed by language.

Cultural Frame Switching

Cultural frame switching can further complicate the investigation of language effects at the formatting stage because scale anchoring may be affected by the reference frame primed by a language. Such differences in scale anchoring may be reflected in the observed differential response styles across cultures. For example, several studies have reported that East Asians avoid extreme responses (Chen, Lee, & Stevenson, 1995; Chun & Campbell, 1974; Hayashi, 1992; Stening & Everett, 1984; Zax & Takahashi, 1967). Although such differences are often attributed to differential emphasis on conflict avoidance and humbleness, it is unclear whether these differences

are an artifact of self-presentation as a result of language priming culture or true differences in perception, independent of language. Moreover, the extent to which respondents use the range of a presented frequency scale as a frame of reference when answering survey questions is also culture dependent. A study by Ji, Schwarz, and Nisbett (2000) demonstrated that Chinese students were influenced by the range of frequency scales only when asked to report private, unobservable behaviors (e.g., having nightmares, borrowing books from the library). However, no scale effects were found for public behaviors (e.g., being late for class), possibly reflecting the importance of "fitting in" in Asian cultures related to monitoring (and thus having better memory representation of) one's and others' public behaviors. In contrast, consistent with previous research on scale effects (for a review, see Schwarz [1996]), American students relied on the presented response scale frequency range to estimate both private and public behaviors. For surveys of bilingual respondents, such findings suggest that, depending on the cultural identity primed by the language of interview, different estimation strategies may be employed.

Codability

Similar to the effect of language codability on retrieval and judgement, response formatting may also be affected by the availability of a label for a given concept. For example, scales may be used differently by speakers of different languages as a result of different scale label codability; thus, the meaning of the same number on a labeled scale may be affected by what language is used. Taken to an extreme, there are cultures whose languages have terms only for one, two and many (Greenberg, 1978), which further limits the ability of their speakers to make comparisons (Hunt and Agnoli, 1991). At this point, little is known how this may affect the cognitive processes in bilinguals whose other language allows for utilization of the whole numeric scale. It can be speculated, that the ability to make comparisons may remain language dependent.

Response Editing

Respondents sometimes edit their responses before reporting them, reflecting social desirability and self-presentation concerns (Sudman et al., 1996). Gender, age, socioeconomic status, and various survey design characteristics have been found to be correlates of socially desirable responding (for a review,

see DeMaio, 1984). Recent work in cross-cultural research suggests that culture influences social desirability through interpretation based on cultural experiences, and response editing depends on the need to conform with particular social norms (Fu, Lee, Cameron, & Xu, 2001; Lee, Xu, Fu, Cameron & Chen, 2001).

Cultural Frame Switching

The same survey question may be perceived to have different levels of socially desirable content depending on the respondent's cultural identity. For example, maintaining harmony and face-saving are more socially desirable traits in Asian cultures than in the Western world (Triandis, 1995). Similarly, mental health is stigmatized in Arab and Hispanic societies (Bazzoui & Al-Issa, 1966; Chaleby, 1987; Okasha & Lotalif, 1979; Silva de Crane & Spielberger, 1981) to a greater extent than in the United States. For bilingual respondents, this means that, depending on the language of the survey interview and the cultural frame primed by it, such questions might be perceived to have different affective characteristics, and respondents would be likely to edit their answers to match the values of the culture associated with the language. The studies by Triandis et al. (1965) and Marín et al. (1983) presented earlier illustrate this effect. For survey practitioners, such language effects would require thorough advance knowledge of where cultural differences related to questions' affective characteristics are to be expected to determine the language assignment of bilingual respondents or to employ questionnaire design techniques that reduce differentially perceived social desirability or sensitivity across language versions.

Summary

A substantial body of literature in psycholinguistics and cross-cultural psychology suggests that language used in survey interviews can affect every stage of the response formation process, and different mechanisms may simultaneously play a role at each step. As our discussion indicates, depending on language, respondents may answer the same question differently as a result of different question interpretation, different mental representations of the question target, a mismatch between the language of encoding and language of recall, different accessible information at the time of the survey request, differential anchoring of response scales, and differential self-presentation concerns.

Two shortcomings of the presented theoretical framework relate to its application. It is desirable to directly connect the outlined model to published survey research and possibly reinterpret puzzling results in light of the proposed language influences, but the existing cross-cultural survey data do not offer such an opportunity. Thus, the proposed framework remains largely speculative. Next, some of the presented mechanisms are demonstrated through research in settings, tasks, and languages that are very different from common survey tasks and languages in which surveys are typically conducted. At this stage, it is unclear to what extent the outlined mechanisms would be detectable in survey responses collected in mainstream (rather than indigenous) languages or whether they are task and language specific. We believe the merits of this theoretical model are to present possibilities for language influences and to stimulate further discussion and action related to these issues.

References

Bazzoui, W., & Al-Issa, I. (1966). Psychiatry in Iraq. *The British Journal of Psychiatry*, *112*, 827–832.

Berry, J., & Sam, D. (1996). Acculturation and adaptation. In J. Berry, M. Segall, & C. Kagitcibasi (Eds.), *Handbook of cross-cultural psychology, volume 3: Social behavior and applications* (pp. 291–325). Boston, MA: Allyn & Bacon.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conception of time. *Cognitive Psychology*, *43*, 1–22.

Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind* (pp. 61–79). Cambridge, MA: MIT Press.

Bower, G. H. (1981). Mood and memory. *The American Psychologist*, *36*, 129–148.

Bower, G. H., Monteiro, K. P., & Gilligan, S. G. (1978). Emotional mood as a context for learning and recall. *Journal of Verbal Learning and Verbal Behavior*, *17*, 573–585.

Bowerman, M., & Choi, S. (2003). Space under construction: Language-specific spatial categorization in first language acquisition. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in studies of language and thought* (pp. 387–427). Cambridge, MA: MIT Press.

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, *236*, 157–161.

Briley, D., Morris, M., & Simonson, I. (2005). Cultural chameleons: Biculturals, conformity motives, and decision making. *Journal of Consumer Psychology*, *15*(4), 351–62.

Bugelski, B. R. (1977). Imagery and verbal behavior. *Journal of Mental Imagery*, *1*, 39–52.

Cannell, C., Miller, P., & Oksenberg, L. (1981). Research on interviewing techniques. *Sociological Methodology*, *12*, 389–437.

Chaleby, K. (1987). Women in polygamous marriages in out-patient psychiatric services in Kuwait. *International Journal of Family Psychiatry*, *8*(1), 25–34.

Chen, C., Lee, S., & Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological Science*, *6*, 170–175.

Chiu, L. H. (1972). A cross-cultural comparison of cognitive styles in Chinese and American children. *International Journal of Psychology*, *7*, 235–242.

Chun, K. T., & Campbell, J. B. (1974). Extreme response styles in cross-cultural research: A reminder. *Journal of Cross-Cultural Psychology*, *5*, 465–480.

DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner & E. Martin (Eds.), *Surveying subjective phenomena* (p. 2). New York, NY: Russell Sage Foundation.

Eich, J. E., Weingartner, H., Stillman, R. C., & Gillin, J. C. (1975). State-dependent accessibility of retrieval cues in the retention of a categorized list. *Journal of Verbal Learning and Verbal Behavior*, *14*, 408–417.

Fu, G., Lee, K., Cameron, C. A., & Xu, F. (2001). Chinese and Canadian adults' categorization and evaluation of lie and truthtelling about prosocial and antisocial behaviors. *Journal of Cross-Cultural Psychology*, *32*, 720–727.

Geertz, C. (1993). *The interpretation of cultures*. New York, NY: Basic Books.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind* (pp. 195–235). Cambridge, MA: MIT Press.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and under water. *British Journal of Psychology*, *66*, 325–331.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612.

Greenberg, J. H. (1978). *Generalization about numeral systems. Universals of human language. Word structure*. Stanford, CA: Academic Press.

Haberstroh, S., Oyserman, D., Schwarz, N., & Kühnen, U. (2002). Is the interdependent self more sensitive to question context than the independent self? Self-construal and the observation of conversational norms. *Journal of Experimental Social Psychology*, *38*, 323–329.

Hayashi, E. (1992). Belief systems, the way of thinking, and sentiments of five nations. *Behaviormetrika*, *19*, 127–170.

Hoffman, C., Lau, I., & Johnson, D. R. (1986). The linguistic relativity of person cognition: An English–Chinese comparison. *Journal of Personality and Social Psychology*, *51*, 1097–1105.

Hong, Y. Y., Chiu, C. Y., & Kung, T. M. (1997). Bringing culture out in front: Effects of cultural meaning system activation on social cognition. In K. Leung, Y. Kashima, U. Kim, & S. Yamaguchi (Eds.), *Progress in Asian social psychology* (pp. 135–146). Singapore: Wiley.

Hong, Y., Morris, M., Chiu, C. Y., & Benet-Martinez, V. (2000). Multicultural minds: A dynamic constructivist approach to culture and cognition. *American Psychologist*, *55*(7), 709–720.

Hunt, E., & Agnoli, F. (1991). The Whorfian hypothesis: A cognitive psychology perspective. *Psychological Review*, *98*, 377–389.

Ji, L., Schwarz, N., & Nisbett, R. E. (2000). Culture, autobiographical memory, and behavioral frequency reports: Measurement issues in cross-cultural studies. *Personality and Social Psychology Bulletin*, *26*, 586–594.

Ji, L., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology*, *87*(1), 57–65.

Kashima, Y. (2000). Conceptions of culture and person for psychology. *Journal of Cross-Cultural Psychology*, *31*, 14–32.

Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist*, *86*, 65–79.

Lee, K., Xu, F., Fu, G., Cameron, C. C., & Chen, S. (2001). Taiwan and Mainland Chinese and Canadian children's categorization and evaluation of lie- and truth-telling: A modesty effect. *British Journal of Developmental Psychology*, *19*, 525–542.

Levinson, S. (1996). Frames of reference and Molyneux's question: Cross-linguistic evidence. In P. Bloom, M. Peterson, L. Nadel, & M. Garrett (Eds.), *Language and space*. Cambridge, MA: MIT Press.

Levinson, S. (2003). *Space in language and cognition: Explorations of cognitive diversity*. Cambridge, UK: Cambridge University Press.

Loftus, E. F., Smith, K. D., Klinger, M. R., & Fiedler, J. (1992). Memory and mismemory for health events. In J. M. Tanur (Ed.). *Questions about questions: Inquiries into the cognitive basis of surveys* (pp. 102–137). New York, NY: Russell Sage Foundation.

Lucy, J. A. (1992). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.

Lucy, J. A., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 257–283). Cambridge, UK: Cambridge University Press.

Lucy, J. A., & Shweder, R. A. (1979). Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist*, *81*, 581–615.

Lucy, J. A., & Wertsch, J. V. (Eds.), (1987). *Vygotsky and Whorf: A comparative analysis. Social and functional approaches to language and thought*. Orlando, FL: Academic Press.

Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, *8*(3), 108–114.

Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology General*, *129*(3), 361–368.

Marín, G., Triandis, H. C., Kashima, Y., & Betancourt, H. (1983). Ethnic affirmation versus social desirability: Explaining discrepancies in bilinguals' responses to a questionnaire. *Journal of Cross-Cultural Psychology*, *14*(2), 173–186.

Means, B., Nigam, A., Zarrow, M., Loftus, E. F., & Donaldson, M. S. (1989). Autobiographical memory for health-related events. *Vital and health statistics*. Washington, DC: National Center for Health Statistics.

Mendoza-Denton, R., Shoda, Y., Ayduk, O. N., & Mischel, W. (2000). Applying cognitive-affective processing system (CAPS) theory to cultural differences in social behavior. In D. L. Dinne, D. K. Forgays, S. A. Hayes, & W. J. Lonner (Eds.), *Merging past, present, and future in cross-cultural psychology* (pp. 205–217). Lisse, Netherlands: Swetz and Zeitlinger.

Nisbett, R. (2003). *The geography of thought. How Asians and Westerners think differently… and why.* New York, NY: Free Press.

Okasha, A., & Lotalif, F. (1979). Attempted suicide: An Egyptian investigation. *Acta Psychiatrica Scandinavica*, *60*, 69–75.

Oyserman, D., Coon, H., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*(1), 3–72.

Oyserman, D., & Lee, S. W. (2007). Priming "culture": Culture as a situated cognition. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology*. New York, NY: Guilford Press.

Pederson, E., E. Danziger, Wilkins, D., Levinson, S., Kita, S., & Senft, G. (1998). Semantic typology and spatial conceptualization. *Language in Society*, *74*, 557–589.

Peytcheva, E. (2019). Can the language of a survey interview influence respondent answers? In T. P. Johnson, B.-E. Pennell, I. A. L. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods*: *Multinational, multiregional, and multicultural contexts (3MC)* (pp. 325–337). Hoboken, NJ: John Wiley & Sons.

Ross, M., Xun, W. Q. E., & Wilson, A. E. (2002). Language and the bicultural self. *Personality and Social Psychology Bulletin*, *28*(8), 1040–1050.

Schrauf, R. W. (2000). Bilingual autobiographical memory: Experimental studies and clinical cases. *Culture and Psychology*, *6*(4), 387–417.

Schrauf, R. W., & Rubin, D. C. (1998). Bilingual autobiographical memory in older adult immigrants: A test of cognitive explanations of the reminiscence bump and the linguistic encoding of memories. *Journal of Memory and Language*, *39*, 437–457.

Schrauf, R. W., & Rubin, D. C. (2000). Internal languages of retrieval: The bilingual encoding of memories for the personal past. *Memory & Cognition*, *28*, 616–623.

Schwarz, N. (1990). Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 98–119). Newbury Park, CA: Sage.

Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Hillsdale, NJ: Erlbaum.

Silva de Crane, R., & Spielberger, C. D. (1981). Attitudes of Hispanic, Black, and Caucasian university students toward mental illness. *Hispanic Journal of Behavioral Sciences*, *3*, 241–255.

Sloutsky, V. M. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences*, *7*, 246–251.

Smith, A. F., & Jobe, J. B. (1994). Validity of reports of long-term dietary memories: Data and a model. In N. Schwarz & S. Sudman (Eds.), *Autobiographical memory and the validity of retrospective reports* (pp. 121–140). Berlin: Springer-Verlag.

Smith, S. M. (1988). Environmental context-dependent memory. In G. M. Davies & D. M. Thomson (Eds.), *Memory in context: Context in memory* (pp. 13–34). Chichester, England: Wiley.

Stening, B. W., & Everett, J. E. (1984). Response styles in cross-cultural managerial study. *The Journal of Social Psychology*, *122*, 151–156.

Strube, G. (1987). Answering survey questions: The role of memory. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 86–101). New York, NY: Springer-Verlag.

Sudman, S., Bradburn, N., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco, CA: Jossey-Bass.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, UK: Cambridge University Press.

Trafimow, D., Silverman, E. S., Fan, R. M.-T., & Law, J. S. F. (1997). The effect of language and priming on the relative accessibility of the private self and the collective self. *Journal of Cross-Cultural Psychology*, *28*, 107–123.

Triandis, H. C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.

Triandis, H. C., Davis, E. E., Vassiliou, V., & Nassiakou, M. (1965). *Some methodological problems concerning research on negotiations between monolinguals*. Urbana, IL: University of Illinois Urbana Group Effectiveness Research Lab.

Tulving, E., & Thompson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.

Wassmann, J., & Dasen, P. R. (1998). Balinese spatial orientation: Some empirical evidence of moderate linguistic relativity. *Journal of the Royal Anthropological Institute*, *4*, 689–711.

Whitney, P. (1998). *The psychology of language*. Boston, MA: Houghton Mifflin.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. New York, NY: Wiley.

Williams, M. D., & Hollan, J. D. (1981). The process of retrieval from very long-term memory. *Cognitive Science*, *5*, 87–119.

Zax, M., & Takahashi, S. (1967). Cultural influences on response style: Comparison of Japanese and American college students. *The Journal of Social Psychology*, *71*, 3–10.